(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 211 047 B1**

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **02.02.94** (51) Int. Cl.5: **C12P 21/02**

(21) Application number: **86900983.7**

(22) Date of filing: **24.01.86**

(86) International application number:
PCT/US86/00131

(87) International publication number:
WO 86/04356 (31.07.86 86/17)

Divisional application 91112901.3 filed on 24/01/86.

(54) ARAB PROMOTERS AND METHOD OF PRODUCING POLYPEPTIDES, INCLUDINN CECROPINS, BY MICROBIOLOGICAL TECHNIQUES.

(30) Priority: **28.01.85 US 695309**
13.11.85 US 797472

(43) Date of publication of application:
25.02.87 Bulletin 87/09

(45) Publication of the grant of the patent:
02.02.94 Bulletin 94/05

(84) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(56) References cited:
EP-A- 0 139 501
US-A- 4 335 104
US-A- 4 520 016

PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 82, 1985; HOFSTEN et al., pp. 2240-2243&NUM;

(73) Proprietor: **Xoma Corporation**
2910 7th Street
Berkeley California 94718(US)

(72) Inventor: **LAI, Jiunu**
10719 Ashby Avenue
Los Angeles, CA 90064(US)
Inventor: **LEE, Jar-How**
2840 Glendon Avenue
Los Angeles, CA 90064(US)
Inventor: **LIN, Yun-Long**
11446 Clarkson Road
Los Angeles, CA 90064(US)
Inventor: **RAY, Dan**
4610 Lindley Avenue
Encino, CA 91316(US)
Inventor: **WILCOX, Gary**
3637 Seahorn Drive
Malibu, CA 90265(US)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

Rank Xerox (UK) Business Services
(3.10/3.09/3.3.3)

PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 80, 1983; ANDREUet al., pp. 6475-6479&NUM;

NATURE, vol. 292 1981; STEINER et al., pp. 246-248&NUM;

EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 127, 1982; QU et al., pp. 219-224&NUM;

CURRENT TOPICS IN MICROBIOLOGY & IMMUNOLOGY, vol. 94/95, 1981, Springer-Verlag, 1981; BOMAN et al., pp. 75-91&NUM;

GENE, vol. 34, 1985; LIN et al., pp. 111-122&NUM;

GENE, vol. 34, 1985; JOHNSTON et al., pp. 137-145&NUM;

JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 260, 1985; KUHN et al., pp. 15907-15913&NUM;

JOURNAL OF BACTERIOLOGY, vol. 156, 1983; MIYADA et al., pp. 765-772&NUM;

PROMOTERS, STRUCTURE AND FUNCTION, 1982; WILCOX et al., pp. 183-194&NUM;

GENE, vol. 14, 1981; HORWITZ et al., pp. 309-319&NUM;

JOURNAL OF BACTERIOLOGY, vol. 143, no. 3, September 1989; J.H. LEE et al., pp. 1325-1331&NUM;

GENETICS, 1983, Pub. J. Wright & Sons (UK); J. FINCHAM, p. 371&NUM;

(74) Representative: **Wright, Simon Mark et al**
Kilburn & Strode
30 John Street
London WC1N 2DD (GB)

Description

Technical Field

The present invention relates to the field of recombinant DNA technology and to the use of araB promoters in the expression of heterologous genes in transformed hosts. This patent also relates to the design, cloning and expression of genes coding for the bactericidal peptide cecropin and analogues thereof.

Background Art

Genetic information, encoded in DNA molecules, is expressed by a series of steps involving transcription of the DNA into mRNA and the subsequent translation of the mRNA into polypeptides or proteins. The expression of the encoded information to form polypeptides is initiated at the promoter site, a region on the DNA molecule to which RNA polymerase binds and initiates transcription. Promoters that have been used in recombinant DNA methods for expressing heterologous genes include the beta-lactamase (penicillinase) and lactose (beta-galactosidase) promoter systems (Chang et al., Nature, 275: 615 (1978); Itakura et al., Science, 198: 1056 (1977); Goeddel et al., Nature, 281: 544 (1979)) and tryptophan (trp) promoter system (Goeddel et al., Nucleic Acids Res., 8: 4057 (1980); EPO Application Publication No. 0036776). Other known promoters include the bacteriophage lambda promoters, (P_L) and (P_R), hut, colicin E_1, galactose, alkaline phosphatase, xylose A, and tac.

The araB gene and its promoter (araB) are located in the L-arabinose operon. The L-arabinose operon (araBAD) in Escherichia coli and in Salmonella typhimurium has been studied. Of particular interest is the L-arabinose operon in S. typhimurium; its sequence is described in Horwitz, A. et al., "DNA Sequence of the araBAD-araC controlling Region in Salmonella typhimurium LT2," Gene, 14: 309-319 (1981), Lin, H.-C. et al., "The araBAD Operon of Samonella Typhimurium LT2, I. Nucleotide Sequence of araB and Primary Structure of Its Product, Ribulokinase," Gene, 34: 111-122 (1985); II. "Nucleotide Sequence of araA and Primary Structure of Its Product, L-Arabinose Isomerase," Gene, 34: 123-128 (1985); III. "Nucleotide Sequence of araD and Its Flanking Regions, and Primary Structure of its Product, L-Ribulose-5- Phosphate-4-Epimerase," Gene, 34: 129-134 (1985). The araBAD operon contains three structural genes which are responsible for the initial metabolism of L-arabinose. Lee, Jar-How et al., "Genetic Characterization or Salmonella typhimurium LT2 ara Mutations," J. of Bacteriology, 158: 344-46 (1984). L-arabinose is first converted into L-ribulose by the araA gene product, L-arabinose isomerase. L-ribulose is then phosphorylated to L-ribulose-5-phosphate by the araB gene product, ribulokinase. The araD gene product, L-ribulose-5-phosphate 4-epimerase, catalyzes the Conversion of L-ribulose-5-phosphate to D-xylose-5- phosphate which then enters the pentose phosphate pathway. The araBAD operon is coordinately Controlled by the inducer L-arabinose and the araC regulatory gene product.

Since, in one embodiment of the invention disclosed and claimed herein, the araB promoter is operably linked to the gene coding for cecropin and inserted into a suitable host for expression of the cecropin protein, it is worthwhile to review background references on cecropins.

The immune system of the Cecropia moth and several lepidopteran insects is characterized by an effective humoral response which is mainly associated with the cecropins, a recently discovered family of antibacterial peptides (Boman, H. G. and Steiner, H., Current Topics In Microbiology And Immunology, 94/95: 75-91 (1981)).

H. Steiner *et al*, Nature **292**: 246-248 (16th July, 1991) describes the purification and characterisation of two proteins obtained from the cecropia moth *hyalophora cecropia*. These proteins have antibacterial activity and the name cecropins A and B are proposed. The tentative amino acid sequences for these two cecropins are given on page 247.

EP-A-0139501 describes the expression of thaumatin in *E. coli* MC1061 cells that have been transformed with the plasmid pING 307 containing the araB gene.

Thus according to a first aspect of the present invention there is provided a replicable expression vehicle comprising a polynucleotide molecule expressible in a given host which comprises the sequence of the araB promoter operably linked to a gene which is heterologous to the host, wherein the gene codes for a polypeptide that is biologically active or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide, where the biologically active polypeptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not thaumatin.

A second aspect of the present invention relates to a process for producing a biologically active heterologous polypeptide in a bacterium transformed to express said peptide comprising:

(a) inserting a polynucleotide molecule comprising the araB promoter operably linked to a heterologous gene which codes for the biologically active polypeptide or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide, where the biologically active peptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not thaumatin, into a replicable expression vehicle;

(b) transforming a bacterium capable of expressing said polypeptide with said expression vehicle containing said polynucleotide molecule;

(c) culturing said transformed bacterium under culturing conditions; and

(d) recovering said polypeptide from said culture.

A third aspect encompasses a process for producing a biologically active heterologous polypeptide by culturing a bacterium transformed with a polynucleotide molecule comprising the araB promoter operably linked to the gene heterologous to the bacterium and coding for the polypeptide, the process comprising the steps of:

(a) incubating a transformed bacterium in a culture medium under selected growth conditions until late exponential growth phase is reached, said bacterium being transformed with an expression vehicle of the first aspect;

(b) transferring the culture of step (a) to fresh culture medium and incubating the culture for a period of time sufficient to allow further growth;

(c) inoculating the culture of step (b) with production medium and incubating the culture at fermentation conditions to achieve maximal bacterial growth and for a time sufficient to achieve a selected cell density;

(d) adding an amount of L-arabinose effective to produce substantial levels of the heterologous polypeptide in the medium; and

(e) recovering the heterologous polypeptide.

In addition, the invention encompasses hosts transformed with the replicable expression vehicle of the first aspect, which is preferably *E. coli, S. typhimurium* or *Serratia marcescens*.

It has now been found that the araB promoter can be used as a promoter in recombinant DNA molecules such that the araB promoter is operably linked to a heterologous gene that codes for a biologically active product. Using the araB promoter to control the expression of heterologous genes has many attendant advantages. The araB promoter control system is tightly regulated. The araB promoter is inducible with L-arabinose; e.g., produced polypeptides are not synthesized prior to addition of L-arabinose to the culture media. Thus, in the absence of L-arabinose there is no expression of the heterologous gene to form the polypeptides. Upon induction with L-arabinose, the polypeptide is transcribed as part of a messenger RNA which initiates at the araB promoter. Once induced, polypeptides are produced quickly and efficiently. Furthermore, the fermentation period is brief as compared with other systems. Importantly, the extent of expression is increased, i.e., the level of production of the heterologous polypeptide is greatly improved, thus making it desirable for commercial use. Finally, the expressed fusion polypeptides form inclusion bodies within the host cell that remain very stable, regardless of increased size, and are amenable to purification for heterologous protein.

The present invention thus provides a polynucleotide molecule expressible in a given host comprising the sequence of the araB promoter operably linked to a gene which is heterologous to said host. The heterologous gene codes for polypeptide(s) that is/are biologically active. The invention also provides for vehicles capable of replication and expression comprising said polynucleotide molecule; hosts transformed with said vehicle; and, fermentation methods for cultivating said hosts for ultimate expression and recovery of the heterologous peptide(s).

The polynucleotide may be a DNA molecule, for example a double-stranded DNA molecule. The vehicle may be a plasmid or a bacteriophage; if it is a plasmid it may also contain a phenotypic selection marker such as antibiotic resistance.

The present invention also provides a process, and tools for use of the process, for the production of cecropin peptides and analogues thereof by recombinant DNA methodology. The cecropins can be expressed by using the aforementioned araB promoter or any other promoter.

The invention may provide a genetic construct capable of expression in a cecropin-sensitive host which comprises a genetic sequence coding for cecropin operably linked to an inducible promoter sequence.

The invention also provides vehicles capable of replication and expression comprising the aforementioned genetic constructs, hosts transformed with said vehicles, as well as methods of producing cecropins using the aforementioned vehicles, constructs and hosts.

The invention further provides a fusion protein of a cecropin with a polypeptide wherein said fusion protein has decreased bactericidal effects toward a given bacterial host.

The fusion protein may carry between both segments a selectively cleavable site.

Preferably, the genetic sequences coding for a cecropin are designed by computational methodology so as to optimize their acceptability by E. coli as expression hosts. Design optimization is also carried out to minimize self-complementarity, to avoid or create restriction endonuclease sites within or outside the sequence and thus facilitate insertion, and to minimize complementarity with the desired expression vehicles. By subjecting the polynucleotide sequence to such optimizations, the invention provides synthetic sequences which, in most instances, are structurally different from those in the natural genes of the various species used as sources for cecropin.

BRIEF DESCRIPTION OF THE DRAWING

Figure 1 shows the amino acid and nucleotide sequences of cecropin A and the synthetic gene used for expression in E. coli.

Figure 2 shows the construction of plasmid pING 1 from plasmid pMH6 and M13mp9. pING carries both the araB and araC genes. The Figure also shows the construction of plasmids pCA1A, pCA1B and pCA1' from plasmid pING1 (as the source of the araB and araC genes).

Figure 3 shows the construction of plasmids pCA3', PCA3A, PCA3B and PCA2' and PCA2A from pMH6 and pCA1A, pCA1B and pCA1'. The differences between pCA1 (or pCA1A and 1B), pCA2' (or pCA2A) and pCA3' (or pCA3A and pCA3B) lie in the varying length between the cecropin A gene and the BamHI site between both ara genes, being 500 bp in pCA1' (or pCA1A and 1B), 1253 bp in pCA2' (or pCA2A) and 1550 bp in pCA3' (or pCA3A and 3B).

Figure 4 shows the construction of plasmid p19C.

Figure 5 shows the construction of plasmid pCA3A- -1 from pCA3A. The Figure also shows the construction of plasmid pCA3D from pCA3A- -1 and p19C (Figure 4).

Figure 6 shows the amino acid and nucleotide sequences of cecropin A, a second chemically synthesized gene used for expression in E. coli.

Figure 7 shows the amino acid and nucleotide sequences of wild type and mutant cecropin A where

(1) denotes CA wild type (pCA3D);

(2) denotes CA mutant A (pCA3A); and

(3) denotes CA mutant B (pCA3B);

↑ denotes the position of mutation.

Figure 8 Shows the construction of the plasmid phTGF5 containing an araB-hTGF fusion.

Figure 9 shows the construction of the plasmid pTGF58.

Figure 10 shows the construction of two plasmids, p115 and p318, carrying a human synthetic calcitonin gene.

Figure 11 shows the construction of a plasmid, pHCT1, containing the gene for the fused polypeptide ribulokinase-human calcitonin (hCT).

Figure 12 shows the construction of plasmid pING1, used as a vehicle for cloning and expression of calcitonin-peptide genes; pING1 carries araB and araC genes.

DEFINITIONS

In the description that follows, a number of terms used in recombinant DNA technology are extensively utilized. In order to provide a clearer and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

Operon. A gene comprising structural gene(s) for polypeptide expression and the control region which regulates that expression.

Operator. A DNA sequence capable of interacting with a specific repressor, thereby controlling the functioning of adjacent gene(s).

Promoter. A DNA sequence within the control region at which RNA polymerase binds and initiates transcription of an adjacent gene(s).

Activator. A protein required for initiation of RNA synthesis by RNA polymerase.

Initiator. A DNA sequence with which an activator interacts to control adjacent genes.

Polynucleotide molecule. A linear sequence of nucleotides linked together by a backbone consisting of an alternating series of sugar and phosphate residues and as used herein can include DNA and RNA

polymers.

Structural gene. A DNA sequence which encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide.

Heterologous gene. A gene that is foreign, i.e. originating from a donor different from the host or a chemically synthesized gene and can include a donor of a different species from the host. The gene codes for polypeptides ordinarily not produced by the organism susceptible to transformation by the expression vehicle.

Biologically active. As used herein means the quality or process of accomplishing an intended effect occurring in a biological system.

Operably linked - As used herein means that the promoter controls the initiation of the expression of the polypeptide encoded by the heterologous gene.

Expression. Expression is the process by which a structural gene produces a polypeptide. It involves transcription of the gene into messenger RNA (mRNA) and the translation of such mRNA into polypeptide-(s).

Cloning vehicle. A plasmid or phage DNA or other DNA sequences which are able to replicate in a host cell, which are characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the DNA, and which contain a phenotypic selection marker suit-able for use in the identification of transformed cells. Markers, for example, are tetracycline resistance or ampicillin resistance. The word "vector" is sometimes used for cloning vehicle.

Expression control sequence. A sequence of nucleotides that controls or regulates expression of structural genes when operably linked to those genes. They include the lac system, the trp system, major operator and promoter regions of phage lambda, the control region of fd coat protein, and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells.

Expression vehicle. A vehicle similar to a cloning vehicle but which is capable of expressing a given structural gene in a host, normally under control of certain regulatory sequences.

Host. Any organism that is the recipient of a replicable expression vehicle, including bacteria and yeast.

Cecropin. This term as used throughout the specification and claims is meant to include a polypeptide from any insect species, which has bactericidal activity in an in vivo or in vitro system acceptable by the art to measure such activity.

The term cecropin is also used in this invention to include any analogue, homologue, mutant, isomer or derivative of a naturally-occurring cecropin, and which shows bactericidal activity in an appropriate system. The term is also meant to include fragments having less than the naturally-occurring number of amino acids, such as partial fragments of natural cecropins or their analogues. The term is also used to include any product which comprises the sequence of a naturally-occurring cecropin or analogue thereof, together with one or more flanking amino acids, preferably at the carboxy terminus, which show cecropin-like bactericidal activity. The term is also meant to include cecropins having less than the number of naturally-occurring amino acids but which still show bactericidal activity.

By "bactericidal activity" is meant to include activity as defined previously, which can either be greater or less than that of naturally-occurring cecropin species. In particular are included cecropin peptides having bactericidal activity which ranges from about 1% of the naturally-occurring species to activities which may be substantially higher (e.g., 10-fold, 100-fold, or higher) than those of the naturally-occurring cecropins.

BEST MODES FOR CARRYING OUT THE INVENTION

1. The araB Promoter

The present invention provides for a polynucleotide molecule expressible in a given host comprising the sequence of the araB promoter operably linked to a gene which is heterologous to said host. The heterologous gene encodes for polypeptide(s) that are biologically active. The invention also provides for associated constructs with the araB promoter. The araB promoter is an inducible promoter, i.e., the expression of the heterologous gene to form the encoded polypeptide is initiated or induced by the addition of L-arabinose. L-arabinose interacts with an activator, the product of the araC gene, to regulate expression of araB. This is a positive control system in contrast to a negative control system; for example, lac, trp, lamda P_L, and tac. Without the L-arabinose addition, the heterologous polypeptide will not be expressed or synthesized. Once induced, the biologically active polypeptide product(s) are produced quickly and efficiently. These polypeptide product(s) typically form as inclusion bodies within the host cells. The inclusion bodies enlarge rapidly as the cell density increases. The inclusion bodies remain very stable

within the host cells. This characteristic results in a reproducible high yield and renders the fermentation monitoring an easy task. To terminate an on-going fermentation run, an accurate decision can be made simply based on the size of inclusion bodies and saturated growth.

The araB promoter in S. typhimurium has the following nucleotide sequence:

```
          +30        +20        +10        +1         -10
5'-CAATTGCCATCGTCTTACTCCATCCAGAAAAACAGGTATGGAGAAACAGTA

     -20        -30        -40        -50        -60
GAGAGTCGCGGCAAAAACCGTCAGGCAGGATCCGCTAATCTTATGGACAAAAAT

-70        -80        -90        -100       -110       -120
GCTAATGCTTTGCAAAGTGTGACGCTGTGCAAATATTCAATGTGGACATTCCAG

     -130       -140       -150       -160       -170
CCATAGTTATAGACACTTCTGTTACTTAATTTTATCGCCTGAACTGTACGCTTT

 -180       -190       -200       -210       -220       -230
TGTTACAAAGCGCTTTTCACAAGCGGGGTTGATACGTGCTTTCATCAAGCGCAA

       -240       -250       -260       -270       -280
AGTCTTGCGGAGACGGAAGCTCTGTCGTCCTGGTCGATATGGACAATTTGTTT

     -290       -300       -310
CTTCTCTGAACATCGGGGGGTAGAGAAATC-3'
```

The araB promoter as part or the araB gene is located in the L-arabinose operon. The L-arabinose operon (araBAD) has been isolated in E. coli and S. typhimurium. However, the practice of this invention is not limited to these two sources for the araB promoter. Other sources can include any organism which contains genes coding for the L-arabinose operon. These sources for the araB promoter can include the genera Pseudomonas, Citrobacter, Xanthomonas, and Erwinia.

In addition, the araB promoter may be synthesized; e.g., by manipulation in the laboratory, rather than of natural origin. In other words, the concept connotes an artificially compounded or even an artificially degraded product. Thus, even though a given complete sequence for a naturally-occurring araB promoter may exist integrated into the genomic DNA or a given organism species, the isolated sequence corresponding to the araB promoter separated from the genomic DNA of the organism, with or without adjacent sequences corresponding to leading polypeptides, start or stop signals, is not naturally occurring and would be considered to be "synthetic." In addition, given the degeneracy of the genetic code, the knowledge of an amino acid sequence does not necessarily and irrevocably lead to the naturally occurring genetic sequence coding therefor.

Optimization of any synthetic sequence comprises four or more possible steps and is applicable to the synthesis of the araB promoter, as well as to synthesis of peptide sequences, including heterologous polypeptides. First, for any given desired sequence, a list of possible DNA codons for each amino acid in such sequence is generated, with those codons ranked in order according to the frequency with which they are used in bacteria or yeast. A preliminary ordering of the codons might be based on the paper by Bennetzen and Hall, Journal of Biological Chemistry, 257: 3026-3031 (1982) (yeast) or Fiers, W., Nature, 260 500 (1976) (bacteria), herein incorporated by reference. A further refinement of the sequence beyond the techniques of these papers is also recommended.

A second factor which influences the choice of any particular codon is the presence or absence of certain restriction enzyme sites that might be used in the process of cloning the gene, so that the use of those enzymes during the cloning process would be facilitated. Optimization of this factor comprises comparing the known endonuclease site sequences (comprising four to six nucleotides per site) with the primary "host-preferred" sequence. A list of restriction endonuclease sites can be found, for example, in Roberts, R. J., Nucleic Acid Research, 11: 1 (1983), r135-r137.

A third factor which influences the choice of particular codons is the need to minimize the internal secondary structure of the synthesized DNA fragments, to prevent them from folding upon themselves and inhibiting the annealing reactions to adjacent DNA fragments. This factor comprises searching a given design sequence so as to avoid undue complementarity of segments thereof, one with another, save for segments adjacent to one another in the intended gene. A search for complementarity (i.e., avoidance thereof) can also be carried out between the designed gene sequence and proposed replication vehicles, such as plasmids or phages.

A fourth factor involved in the optimization is the avoidance of sequences rich in AT base pairs (about 5 or more), especially when preceded by a sequence rich in GC base pairs, to avoid premature termination of transcription.

A fifth factor influencing the choice of codons is the avoidance of RNase sites so that the message is stable.

Finally, if any of two possible codons could be used, it is preferred to utilize that which will maximize expression in microbial genomes (see, for example, Fiers, et al., Nature, 260 500 (1976); Grosjean, et al., Gene, 18: 199-209 (1982); and Riggs U.S. Patent 4,366,246, column 6, all of which are herein incorporated by reference).

Most preferably, a computer program designed to carry out the necessary comparisons, to optimize expression in bacterial microbes or yeast, is utilized for the optimization.

In one embodiment of this invention, the inducible araB promoter is operably linked to a genetic sequence coding for a polypeptide that is biologically active, and the resulting genetic construct is introduced into or forms part of an expression vehicle. The expression vehicle is then utilized to transform an appropriate host. The host is fermented under selected culturing conditions to achieve optimum growth. The araB promoter is not active until treated with L-arabinose, which induces the promoter to initiate expression of the heterologous gene. The sequence of actions include transcription of the gene into mRNA, and the translation thereof into the polypeptide product(s).

In another embodiment of this invention, the araB promoter is operably linked to a genetic sequence coding for a heterologous polypeptide that is biologically active, and this genetic sequence is operably linked to a second genetic sequence coding for another polypeptide. The expression yields a fusion or precursor protein comprising both the amino acid sequence of the second polypeptide and that of the desired heterologous polypeptide, and containing a selective cleavage site adjacent to the desired amino acid sequence.

The cleavage site is preferably methionine, although the site may be any preferred site known in the art. The desired heterologous polypeptide should preferably lack internal cleavage sites corresponding to the actual selected cleavage site. Other known cleavage sites include Asn-Gly, Asp-Pro, Lys, Arg, and Lys-Arg.

Selective cleavage of the fusion or precursor protein is typically effected outside of the replicative environment of the expression vehicle. In this post-translational step, the fusion or precursor protein is clipped by selective treatment. For example, when methionine is the cleavage site, the fusion or precursor protein is treated with cyanogen bromide to clip the desired heterologous polypeptide. With other known cleavage sites, the clipping treatment includes hydroxylamine, acid, trypsin, and Lys-Arg cleavage enzyme.

Methods for preparing fused, operably linked genes and expressing the same in bacteria are known and are shown, for example, in U.S. 4,366,246, herein incorporated by reference.

The genetic constructs and the methods involved herein can be utilized for expression of the heterologous polypeptides in bacterial hosts.

For example, E. coli K12 strain 294 (ATCC 31446) and strain MC1061 (ATCC 39450) are particularly useful. Other microbial strains which may be used include, but are not limited to, E. coli X1776 (ATCC 31537). The aforementioned strains, as well as E. coli W3110 (F^-, $lambda^-$, prototrophic (ATCC 27325)), and other enterobacteriaceae such as S. typhimurium or Serratia marcescens, and various pseudomonad species may be used.

In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as specific geries which are capable of providing phenotypic selection in trans-

formed cells. For example, E. coli is typically transformed using pBR322, a plasmid derived from an E. coli species (Bolivar, et al., Gene, 2: 95 (1977)). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmids must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

The araB promoter may be operably linked to a genetic sequence coding for any heterologous polypeptide or protein that is biologically active. Examples of such polypeptides or proteins include, but are not limited to, enzymes, hormones, hemoglobin, antibodies, structural proteins, alpha, beta and gamma interferons, interleukins, insulin, and tissue plasminogen activators. Specific examples include human tumor growth factor, calcitonin, and cecropin.

The transformed host can be fermented and cultured according to means known in the art to achieve optimal cell growth. The preferred fermentation and production procedure of this invention, described below, can be used to achieve large scale production of heterologous polypeptides. In addition, by using the araB promoter operably linked to the heterologous gene, the extent of expression of the heterologous gene is increased.

The preferred fermentation procedure is as follows: The transformed host, preferably bacteria, more preferably E. coli, is introduced into a culture medium containing nutrient materials that meet the growth requirements of the bacterium. Such materials may include carbon and nitrogen sources, minerals, amino acids, purines, pyrimidines, and vitamins. The preferred culture medium also comprises a metabolite in an amount sufficient for phenotypic marker resistance, such metabolite being, for example, tetracycline or ampicillin. The host is grown under culturing conditions selected to achieve maximum growth rate. Temperature conditions will depend upon the host, but typically the optimum range is about 30°C to about 40°C, with 37°C being the most preferred for transformed E. coli. Oxygen is also provided to the medium. The host is allowed to grow until late exponential phase and then transferred to fresh culture medium.

The hosts are then inoculated with production culture medium. During this step, the bacteria will continue to divide and grow until the bacteria reach a concentration, saturation density, at which the bacteria no longer divide, but are still viable. The time sufficient to reach the saturation density is dependent on the medium, the genotype of the host, the temperature, and the degree of aeration.

Initially, the fermentation and culturing conditions for the production step are the same as those given above for the culture step, except that as dissolved oxygen is consumed during growth, the agitation and aeration rate are increased accordingly to maintain a minimum dissolved oxygen (D.O.) at 30%. The preferred pH range is between about 6 to 8. The pH of the above-described production medium is consistently self-adjusted at 6.5-6.8, which is optimum for E. coli growth. Thus, acid and base control of the pH medium is unnecessary for the E. coli host; however, it may be necessary to adjust the pH medium for other hosts. When optimal cell density is reached, optimally $OD_{600}{}^{10}$ for E. coli, L-arabinose is added in an amount sufficient to induce the synthesis of the heterologous gene to form the polypeptide. As described above, the polypeptide forms inclusion bodies within the host. The peptide is then recovered according to means known in the art such as filtering or precipitation. If a fusion protein is the resulting expressed heterologous peptide, the fusion protein is recovered and then may be treated in a post-translation step to separate the desired heterologous polypeptide. The heterologous polypeptide can then be recovered and purified according to known means.

The use of the araB promoter will be described, for convenience, relating to various embodiments of the invention comprising cecropin, although it is to be understood that the araB promoter may be operably linked to a genetic sequence coding for any polypeptide or protein that is biologically active.

2. Methods of Microbial Production of Cecropins Using Cecropin-Sensitive Hosts

Part of the present invention also provides methods of microbiological production of cecropins using cecropin-sensitive hosts. One concept of the invention is to express a gene sequence coding for cecropin while, simultaneously, avoiding or delaying the bactericidal effects of the product.

In one embodiment, the genetic sequence coding for cecropin is linked to an inducible promoter, and the resulting genetic construct is introduced into an expression vehicle. The expression vehicle is then utilized to transform an appropriate host and the host is fermented under normal conditions wherein the promoter is not active. After an appropriate period of time, such as, for example, at a time when the cells are in stationary phase, the promoter is induced as by varying an outside environmental factor such as salt concentration, light, presence or absence of a metabolite, a metal, and the like, this change leading to transcription of the cecropin genetic sequence into mRNA, and then translation thereof into bactericidal cecropin. Even though the resulting cecropin is bactericidal and destroys the bacterial host, such destruc-

tion does not occur until late in the fermentation cycle. Examples of regulated promoters are lambda P_L and P_R, lac, gal, trp, ara, hut, and the like.

In the second preferred embodiment, it has been discovered that if a genetic sequence coding for cecropin is operably linked to a polypeptide other than said cecropin, such that the expression yields a fusion or precursor protein comprising both the amino acid sequence of cecropin and that of the additional polypeptide, and containing a selective cleavage site adjacent to the desired cecropin amino acid sequence, the resulting fusion protein is not bactericidal. Bactericidally active cecropin can then be isolated post-translation by selective cleavage.

Most commonly, cleavage will be effected outside the replicative environment of the expression vehicle, such as, for example, following harvest of the microbial culture. Thus, the additional polypeptide robs the cecropin of its bactericidal effect pending extra-cellular cleavage, allowing the survival of the host for long enough time to yield high levels of the desired product. Preferably, the cecropin will lack internal cleavage sites corresponding to the selective cleavage site employed to shed the superfluous polypeptide, although this is not necessarily an absolute condition. Since the cecropins are methionine free, cyanogen bromide cleavage at the methionine adjacent to the cecropin sequence is effective.

Preferably, the genetic sequence coding for the superfluous polypeptide is transcribed in advance of the structural gene of the cecropin, but this need not necessarily be the case, as it may also be possible to express the superfluous polypeptide in a position adjacent to the C-terminal of the cecropin.

The nature of the superfluous adjacent polypeptide is not critical. It could either be part, whole or repetitive units of a known structural, enzymatic, hormonal, or other physiologically relevant proteins. Alternatively, it could be a non-functional polypeptide. Without being bound by any particular theory, the inventors speculate that the increased length of the fusion protein somehow "masks" the bactericidal properties of the cecropin due to the varying conformation of the overall polypeptide. Preferably, the genetic sequence coding for the adjacent superfluous polypeptide should be at least about greater than 300 base pairs in length, preferably between about 400 bp and 5 Kb, most preferably between 400 bp and 2 Kb. This corresponds to a superfluous polypeptide of preferably between about 100-1700 amino acid residues.

Any of a large number of superfluous polypeptides can be fused to the desired cecropin peptide sequence. The polypeptide gene sequence can either be prepared by organic synthesis, in which case optimization procedures would be recommended, or might be prepared by such techniques as reverse transcription of appropriate mRNA. Enzymatic coupling of the gene sequence for the polypeptide to the gene sequence for the structural cecropin peptide would then follow the preparation of the cDNA. Enzymatic coupling can be either by blunt ligation or by the provision of cohesive termini, comprising one of the two strands of a restriction endonuclease recognition site. Examples of superfluous polypeptides are beta-galactosidase or ribulokinase (encoded for by the araB gene). Enzymes and structural proteins are preferred. Other products which can be used are products encoded by the following genes: aceA or aceB, araA, araB, araC, araD, argG, aroB, lacA, serA, purA, trpA, trpB, trpC, trpD, trpE, tyrA, and the like. The superfluous polypeptide is normally free of glycosylation. The araB promoter is the promoted to be used.

In yet another embodiment of the invention, the cecropin genetic sequence is operably linked to the sequence for a superfluous polypeptide capable, in the fusion product, of inhibiting or inactivating the bactericidal activity of cecropin and, in addition, to an inducible promoter. In this manner, both the effect obtainable through the fusion protein technique and the effect obtainable through the use of the inducible promoter can be exploited advantageously and simultaneously.

Although the present invention results in methods for producing cecropins in cecropin-sensitive hosts, e.g., bacterial hosts, the genetic constructs prepared herein and the methods involved can also be utilized for expression of cecropin in other, non-cecropin sensitive hosts. These non-cecropin sensitive hosts include yeasts and mammalian cell cultures. Useful yeast and mammalian hosts and vectors are well known to those of skill in the art, and reference is made, for example, to European Patent Publication 0093619 published November 9, 1983. Bacterial hosts can include those mentioned hereinabove with the araB promoter, as well as bacterial hosts such as the genera Pseudomonas, Citrobacter, Xanthomonas, and Erwinia. Any plasmid vector compatible with these hosts, as described above with the araB promoter, can be used.

Another preferred promoter for cecropin is lambda (P_L). The genetic construct for cecropin and the superfluous polypeptide can be placed under the control of the leftward promoter of bacteriophage lambda (P_L). This promoter is one of the strongest known promoters which can be controlled. Control is exerted by the lambda repressor, and adjacent restriction sites are known. A temperature sensitive allele of CI gene can be placed on the vector that contains the cecropin complete sequence or a different vector. When the temperature is raised to 42°C, the repressor is inactivated, and the promoter will be expressed at its maximum level. The amount of mRNA produced under these conditions should be sufficient to result in a

cell which contains about 10% of its newly synthesized RNA originated from the P_L promoter. In this scheme, it is possible to establish a bank of clones in which a functional cecropin fusion construct sequence is placed adjacent to a ribosome binding sequence, and at varying distances from the lambda P_L promoter. These clones can then be screened and the one giving the highest yield selected.

The expression of a cecropin sequence can also be placed under control of other regulons which may be "homologous" to the organism in its untransformed state. For example, E. coli chromosomal DNA contains a lactose or lac operon which mediates lactose digestion by elaborating the enzyme beta-galactosidase. The lac control elements may be obtained from bacteriophage lambda plac5, which is infective for E. coli. The phage's lac operon can be derived by transduction from the same bacterial species. Regulons suitable for use in the process of the invention can be derived from plasmid DNA native to the organism. The lac promoter-operator system can be induced by IPTG.

Of particular interest in the present invention is to provide synthetic polynucleotide sequences coding for the cecropin peptides. In a preferred embodiment of the present invention, the synthetic sequence of the cecropin peptide (with or without adjacent sequences) is optimized so that expression thereof will be compatible with a variety of hosts such as yeasts and bacteria, especially the latter. In particular, optimization of the expression of any given sequence in E. coli is of great importance. Thus, after performing such optimization procedures as indicated above, the actual genetic sequence of the cecropin peptides with or without adjacent sequences, will usually be distinct from the naturally-occurring sequence in the original species.

In addition, the design of the desired gene for the fused product should preferably incorporate codons for amino acids at the cleavage site as methionine (cleavable by cyanogen bromide), tryptophan (cleavable by o-iodosobenzoic acid), glutamic acid (cleavable by Staph. protease) and the like.

In one embodiment of the invention, any given codon in the desired DNA sequence for the fusion product can be mutagenized at will through site-directed mutagenesis. Thus, it is possible, after synthesis of the desired DNA sequence, to introduce into the sequence a cleavable site. Site-directed mutagenesis is known, and reference is made to Wallace et al., Science, 209: 1396-1400 (1980), herein incorporated by reference.

The amino acid residue or residues following the potential C-terminal residue in the cecropin-peptide may be followed by yet another polypeptide ("trailing" polypeptide) of varying length and of structure similar or different than a leading polypeptide if one is present, as described above. In such case, it is convenient to provide for the same or different selective cleavage sites between the C-terminal amino acid residue and the trailing polypeptide sequence. These cleavable sites may or may not be the same as those present between the leading polypeptide and the structural gene for the cecropin peptide.

In other words, the superfluous polypeptide may be present as a leading peptide, a trailing polypeptide or both.

Where the structural gene of the desired cecropin peptide is to be inserted into a vehicle for expression as such, the gene would be preceded by a "start" codon, and if followed by trailing sequences, one or more termination or stop codons. When the expression product is a fusion protein comprising both the cecropin peptide and part or whole of a polypeptide, the start codon may be placed prior to the N-terminus of the polypeptide if it is leading.

Methods for the syntheses of polynucleotides are well known to one skilled in the art. Reference is made, for example, to the triester method of Itakura et al., Journal of American Chemical Society, 97: 3727 (1975).

The cecropins produced by the methods of the invention can be used as broad anti-microbial agents directed toward specific applications. Such applications include, for example, the use of the cecropins as preservatives in processed meat products (target organisms: (1) Clostridium botulinum, (2) Lactobacilli, (3) Micrococci); anti-caries agents in oral hygiene products (target organism: Streptococcus mutans); agents useful in the treatment of vaginal yeast infections (target organism: Candida albicans); and as anti-bacterial agents in deodorants (target organisms: (1) Micrococci, (2) Diphtheroids). The relative effectiveness of cecropin-like peptides for the applications described can be readily evaluated by one of skill in the art by determining sensitivity of any organism to one of the cecropin peptides. The same bacterial screen used in vitro can be utilized to determine dosages and concentrations.

Having now generally described this invention, the same will be better understood by reference to a specific example which is included herein for purposes of illustration only and is not intended to be limiting unless otherwise specified.

METHODS

Amino Acid Compositions

Hydrolysis of cecropin polypeptides was accomplished using 5.5M HCl (Pierce) at 110°C for 24 hours in vacuo. Samples were dried under vacuum at 40°C and resuspended in 200 µl of Beckman Model 6300 sample buffer (low pH citrate).

Amino acid analysis was performed using a Beckman Model 6300 analyser. Postcolumn ninhydrin was used for detection. Data was recorded using a Hewlet Packard (HP) integrator Model 3390. The individual peaks were recorded as the sum of the signals at 540 and 440 nm. Beckman standards containing each amino acid were used to calibrate the integrator and sperm whale myoglobin (Applied Biosystems) was used to verify the calibration.

Protein Sequencing

An Applied Biosystems Model 470A protein sequenator was used for all protein sequence determinations. Specifications for the sequencing methodology follow the Hunkapiller-Hood program. The system uses a non-vacuum program with cleavage by methanolic-HCl.

PTH Determinations

PTH determinations were accomplished using an HP model 1090A HPLC with an HP model 3390A integrator for data presentation. An IBM cyano-propyl column was used for separation of the PTH amino-acids. The buffer was 30 mM NaHAc pH 5.1 containing 5% tetrahydrofuran. The flow rate was 1 ml./min. and temp. 37°C.

Samples were dissolved after drying in vacuo with 30 µl of 33% acetonitrile in water for 30 min. The PTH standards used were from Pierce Chemical Co.

EXAMPLE 1

Synthesis and Cloning of a Synthetic Cecropin A (CA) Gene

A. Synthesis of the CA Gene

A gene coding for CA was designed by computer to incorporate codons normally found in highly expressed E. coli proteins. At the end of the gene were incorporated 4-bp overhangs to permit the ligation of SalI and EcoRI sites respectively. To facilitate the construction of various sizes of fused proteins a BamHI site was included following the SalI site. Also, by using a computer program, optimization procedures as described previously were taken into account. The gene was divided into eight oligonucleotides ranging from 23 to 37 bases in length (Figure 1).

The eight single-stranded DNA fragments were synthesized according to the solid-phase phosphotriester method as described by Ito, et al., Nucleic Acids Research, 8: 5491 (1982).

Several modifications and improvements were made, including the following:

(1) The coupling reaction was carried out at 37°C for 40 minutes with gentle shaking.

(2) For the deprotection of the DNA after final coupling reaction, the DNA resin (10 mg) was reacted with a 0.5 M 2-pyridinealdoxime-tetramethyl guanidinium solution (2 ml) in pyridine-water (8:2 v/v) for 60 hours at 37°C with shaking. The combined solution was evaporated and treated with NH_4OH (28%, 3 ml) in a sealed ampoule at 60°C for 36 hours. After evaporation of the ammonia, the solution was extracted with ether three times and then evaporated. For the purification of the DNA, the residue was dissolved in 0.1 ml of 50 mM triethylammonium bicarbonate pH 7.5 (TEAB) and applied to a Sephadex G-50 column (1 x 50 cm). One ml fractions were collected. The first few fractions, at the leading edge of the major 260nm-absorbing peak, contained the desired product. These fractions were evaporated, then purified further by high pressure liquid chromatography (HPLC). HPLC purification was carried out on a Bondapak C18^R (Waters) column at 55°C using a linear gradient of acetonitrile (10-40%) in 10mM triethylammonium acetate buffer (pH 7.8). The DMT group was hydrolyzed by treatment with 80% acetic acid (0.1 ml) for 25 minutes at room temperature followed by evaporation and then evaporated with 0.1 ml water to remove the acetic acid completely.

B. CA Gene Assembly by Ligation

The 5'OH termini of the chemically synthesized fragments 1 through 8 were separately phosphorylated in the presence of 10 µl of solution containing:

70mM Tris-HCl (pH 7.6)
10mM $MgCl_2$
5mM dithiothreitol
66 µM gamma-^{32}P-ATP (1 µCi)
400 ng of DNA and 2 units of T4 polynucleotide kinase

The reaction was held at 25°C for 15 minutes then 1 ml of 10mM unlabelled ATP was added to continue the reaction for another 15 minutes. To check the purity, 10% of the phosphorylated DNA was analyzed by the standard method using polyacrylamide gel (15%) electrophoresis in the presence of 7M urea.

Four hundred ng of the phosphorylated DNA fragments 2, 3, 4, 5, 6 and 7 were heated at 95°C for five minutes to inactivate the T4 polynucleotide kinase. Then 400 ng of unphosphorylated fragments 1 and 8 were added and heated at 95°C for another five minutes and then cooled to 25°C slowly (1°C per 10 min.). The eight DNA fragments were ligated in a total volume of 100 µl in the presence of 66mM Tris-HCl (pH 7.5), 6.6mM $MgCl_2$, 10mM dithiothreitol, 0.4mM ATP and 2 units of T4 DNA ligase for two hours at 25°C.

C. Construction of Plasmids pING 1, pCA1A, pCA1B and pCA1'

The construction scheme is shown in Figure 2.

pMH6 is described in Horwitz, A. H. et al., Gene, 14:309-319 (1981), and M13mp9 is commercially available.

1. The CA gene fragments obtained from Step B, supra, were ligated to plasmid pING 1 which had been pretreated with restriction endonuclease SalI and EcoRI, then digested with restriction endonuclease SmaI to decrease the transformants carrying pING 1.
2. The SmaI-treated plasmid was transformed into E. coli MC1061.
3. The colonies which contained plasmids carrying the CA gene fragment were identified by colony hybridization using the synthetic DNA fragment 7 (Figure 1) as a probe. Three independent clones which contained the CA fragment were found in 1,000 colonies. Each of the isolated plasmids was digested with SalI and EcoRI to check the size of the excised fragment. The nucleotide sequence analysis of the CA insert was performed by the dideoxy chain termination procedure of Sanger et al., PNAS, 74: 5463-5467 (1977) with some modification (Wallace et al., Gene, 16: 21-26 (1981)). Two sequences were determined and are shown in Figure 7(2) and 7(3). The plasmids containing these sequences were designated pCA1A and pCA1B (Figure 2). Each of these sequences differed from the wild type CA sequence at the positions indicated by the arrows in Figure 7(2) and 7(3).

Various methods can be used to generate a wild-type sequence, for example in vivo or in vitro recombination of pCA1A and pCA1B or screening of additional independent clones. The plasmid containing the proposed wild type sequence is designated pCA1' (Figure 2).

In the description that follows, plasmids with primes (') denote wild type CA sequences, whereas plasmids without primes are derived from mutants pCA1A or pCA1B.

D. Construction of plasmids pCA2', pCA2A pCA3', pCA3A and pCA3B

This is shown in Figure 3. pCA1 (or pCA1A or pCA1B) is digested with BamHI and, after treatment with polymerase plus dTNP, followed by SstI digestion, yields a fragment (1) which is blunt ended at one end and carries an SstI overhang at the other. This fragment is ligated with a fragment obtained from pMH6 by NarI digestion, blunt end formation and SstI treatment, to give, after T4 ligation, pCA3' (or pCA3 or pCA3B). The fragment (1) is also ligated with a fragment obtained from pMH6 by HgiAI digestion, blunt end formation and SstI treatment, to give, after T4 ligation, pCA2' (or pCA2A).

The desired products are sought by transforming E. coli MC1061 with the ligation products. Plasmids are isolated by the minilysate procedure indicated supra. Each of the isolated plasmids is digested with BamHI and PstI respectively. Plasmids from each group having the correct size and correct orientation of the BamHI fragment are designated pCA2 (or pCA2A) and pCA3' (or pCA3A or pCA3B).

E. Expression of araB-CA fusion protein

E. coli cells containing araB-CA fused genes in plasmids pCA1A, pCA2A, pCA3A, pCA3B and pCA3D (note: pCA3D will hereafter interchangeably be used to denote wild type CA containing plasmid pCA3') were grown at 37°C in media containing 1.5% tryptone, 1.0% yeast extract and 0.5% NaCl (TYE) with ampicillin at 50 ug/ml. At cell densities of $OD_{600} = 0.2$, cultures were treated with L-arabinose to a final concentration of 0.5% to induce expression of araB-CA fusion protein, and were harvested when cell densities reached $OD_{600} = 1.5$ to 1.7, by centrifugation at 4000 RPM at 4°C for 20 minutes in a Beckman JS-4.2 rotor. Cells were washed once by resuspending in one-half original culture volume of 50mM phosphate buffer (pH 6.6). Washed pellets of cells containing plasmids pCA1A, pCA2A, and pCA3A were extracted in one-tenth volume phosphate buffer and 1% sodium dodecyl sulfate (SDS) and analyzed on a 10% polyacrylamide SDS-denaturing gel. By this analysis, it was shown that E. coli cultures containing plasmid pCA3A produced more araB-CA protein than either pCA1A or pCA2A-transformed culture.

F. Fractionation of induced E. coli cells containing plasmids pCA3A, pCA3B and pCA3D

Washed pellets from Step E were resuspended in one-tenth original volume of phosphate buffer, chilled to 0°C, then the cell was broken in a French pressure cell. The subcellular components were then centrifuged at 4000 RPM at 4°C for 20 minutes in the rotor. Analysis of the resulting supernatant and 4000 RPM pellet on a 10% polyacrylamide SDS denaturing gel showed that all of the araB-CA fusion protein of pCA3A, 3B and 3D were found in the 4 K pellet.

G. Purification of the synthetic CA

(a) Cyanogen bromide cleavage

1. The 4 K pellet fraction which contained the araB-CA fused protein obtained from the last step was mixed with 90% formic acid to give 0.7-1.1 mg/ml protein in 70% formic acid. A 10-fold excess of cyanogen bromide (1 gm/ml stock in acetonitrile) by weight was added. The reaction mixture was flushed with nitrogen and incubated at room temperature overnight.
2. The formic acid and cyanogen bromide were removed under a stream of nitrogen. The residue was dissolved in 70% formic acid and dried under a stream of nitrogen twice more.
3. A solution of 0.1% trifluoroacetic acid in water was added; this dissolved the cleaved CA completely at 1 mg protein/ml.

(b) High pressure liquid chromatography of cyanogen bromide fragments

The cyanogen bromide-cleaved fused protein was chromatographed by HPLC to isolate the fragment from the CA portion of the molecule. A C-18 reverse-phase column (Waters) was used. A gradient was run with 0.1% trifluoroacetic acid (TFA) in water (buffer A) and 0.1% TFA in acetonitrile (buffer B). The starting eluent contained 20% buffer B; at 2 minutes after sample injection, a gradient of 20% B to 60% B over 60 minutes was initiated. The flow rate was 1 ml/minute. The elution profile was monitored at 280 nm. Various peaks in the chromatogram of the cyanogen bromide-cleaved fused protein which eluted were collected to test the bactericidal activity as described later below (Example 4).

H. pCA3A Codes for a Mutant Cecropin Polypeptide

Table I

Amino Acid Composition of Cecropin-A Produced by pCA3A in E. coli		
Amino Acid	Mutant A	
	Experimental Values	Theoretical Values*
Asp	2.9	3
Thr	0.1	0
Glu	2.7	2
Pro	2.4	2
Gly	6.3	4
Ala	3.0	3
Val	3.6	4
Ile	2.9	4
Leu	1.2	1
Phe	1.1	1
Lys	6.2	6
Arg	2.1	2
Trp	ND	1
Ser	1.3	1
Cys	0	0
Met	0	0
Tyr	0.1	0
His	0.4	0
N.D. = not determined.		

* Theoretical Values: Obtained from DNA sequencing analysis of each mutant clone.

The nucleic acid and the amino acid sequences are shown in Figure 7.

EXAMPLE 2

Demonstration of Another Mutant Cecropin Polypeptide

The cells of plasmid pCA3B were fractionated, CA sequences were detected, and the synthetic CA was purified and analyzed as shown in Example 1, Parts F-G. Amino acid composition and sequence data, shown in Table II, indicated that the second clone also contained a CA gene sequence coding for a mutant cecropin.

Table II

Amino Acid Composition of Cecropin-A Produced by pCA3B in E. coli		
Amino Acid	Mutant B	
	Experimental Values	Theoretical Values*
Asp	2.3	2
Thr	1.0	1
Glu	4.2	4
Pro	1.0	1
Gly	4.1	3
Ala	5.0	5
Val	3.6	4
Ile	3.9	5
Leu	1.1	1
Phe	0.9	1
Lys	6.6	7
Arg	1.0	1
Trp	ND	1
Ser	1.1	1
Cys	0	0
Met	0	0
Tyr	0	0
His	0.1	0
N.D. = not determined.		

* Theoretical Values: Obtained from DNA sequencing analysis of each mutant clones.

The nucleic acid and amino acid sequences are shown in Figure 7.

EXAMPLE 3

Construction of Plasmid pCA3D Coding for Wild Type Cecropin A

The construction scheme is shown in Figures 4 and 5.

A. Construction of p19C

1. Plasmid pING 1 was digested to completion with endonucleases FokI and then BamHI; the resulting digestion was halted by phenol/chloroform (ratio 1:1) extraction and the aqueous residue was washed with ether and then passed through a Bio-Gel^R P-10 column to remove the phenol/chloroform.

The result of FokI digestion produced a "CTAC" 5' overhang. This 76 base pair BamHI, FokI fragment contains the araB promoter.

2. A new CA gene fragment (see Figure 6) was constructed using the same procedure as described in Example 1, Step B, except that the two DNA fragments No. 1 and 5 were replaced with No. nn-1 and nn-5 (sequence shown in Figure 6) to produce a "GATG" 5' overhang at its N-terminus and an EcoRI site overhanging ATTT at its C-terminus.

3. pBR322 was digested to completion with the endonucleases BamHI and EcoRI. The resulting digestion was halted as described.

4. The BamHI-FokI fragment from Step 1, supra, and the new CA gene from Step 2, supra, were ligated to the larger BamHI-EcoRI fragment of pBR322.

5. The ligated product was digested with endonuclease HindIII to decrease the number of transformants carrying pBR322.

6. The HindIII-treated plasmid was transformed into E. coli MC1061.

7. The colonies which contained plasmids carrying the CA gene fragment were identified by colony hybridization using the synthetic DNA fragment 7 (Figure 6) as probe. Three independent clones which

contained the CA fragment were found in 1,000 colonies. Each of the isolated plasmids was digested with BamHI and EcoRI; a plasmid able to release the correct fragment was designated p19C. The nucleotide sequence analysis of the CA inserts wad performed by the dideoxy chain termination procedure of Sanger et al., PNAS, 74: 5463-6467 (1977), with some modification (Wallace et al., Gene, 16:21-26 (1981)). p19C contained the correct sequence for the CA gene. The CA gene was placed directly downstream of the araB promoter without any araB coding sequence in between.

B. Construction of pCA0

1. The BamHI-EcoRI fragments of p19C were excised by digestion with excess amounts of restriction endonuclease EcoRI and BamHI. The plasmids were also digested with PvuI to decrease the chance that the excised fragments would be religated to the plasmid in the next ligation step.
2. The BamHI-EcoRI fragments from Step A.7 were ligated to plasmid pING 1 which had been pretreated with restriction endonuclease BamHI and EcoRI, then digested with restriction endonuclease SmaI to decrease the transformant carrying pING 1.
3. The SmaI-treated plasmid was transformed into E. coli MC1061.
4. The colonies which contain plasmids carrying the CA gene fragment were identified by plasmid characterization. Each of the isolated plasmids was digested with BamHI and EcoRI to check the size of the excised fragment. One plasmid that had the correct size was designated pCA0. The nucleotide sequence analysis of the CA inserts was performed by the dideoxy chain termination procedure. This pCA0 contained the correct sequence for the CA gene. This pCA0 containing the complete araB regulatory gene and CA gene is placed directly after the araB promoter to express the cecropin-A directly, without creating an araB-CA fusion protein.

C. Construction of Plasmid pCA3A- -1

The purpose of this construction is to delete the AvaI-PvuII region from pCA3A in order to eliminate one XmnI site. The deletion is also able to increase the plasmid copy number in E. coli cells.

1. pCA3A was digested with endonuclease AvaI and PvuII and then filled in with the Klenow fragment of DNA polymerase I in the presence of dNTP to produce a blunt end.
2. The plasmid from Step 1. was religated and then digested with AvaI and PvuII to decrease the number of transformants containing original pCA3A.
3. The AvaI, PvuII-treated plasmid was transformed into E. coli MC1061.
4. Colonies which contained plasmids deleted for the AvaI-PvuII region were identified by colony hydridization using the synthetic DNA fragment 5'-TCATCAGCGTGGTCG-3' as a probe. Forty-six independent clones with the AvaI-PvuII region deleted were found in fifty colonies. Each of the isolated plasmids was digested with XmnI to check the size of the excised fragments. One of the plasmids that had the correct size was designated pCA3A- -1.

D. Final Assembly of Plasmid pCA3D

p19C contains the wild-type cecropin-A sequence but lacks a convenient restriction site at its N-terminus to create a protein fusion with araB.

pCA3A has a deletion of two base pairs occurring near the C-terminus of the CA gene, causing it to produce a mutant araB-CA fused protein.

These two plasmids were recombined to construct a plasmid that had the wild-type CA gene fused to araB and produce an araB-CA fused protein.

The construction scheme is shown in Figure 5.

1. Plasmid p19C (1 µg) was digested to completion with the endonuclease XmnI; the resulting digestion was halted by heating at 65°C for 10 min.
2. Plasmid pCA3A- -1 (0.1 µg) was digested to completion with the endonuclease XmnI, the resulting digestion halted by heating at 65°C for 10 min.
3. DNA from steps D.1 and D.2 were mixed and ligated.
4. The ligated product was digested with PvuII to decrease the transformants carrying p19C.
5. The PvuII-treated plasmid was transformed into E. coli MC1061.
6. The colonies which contained plasmids carrying the wild-type araB-CA gene and able to produce an araB-CA fused protein were identified by DNA sequence analysis, as previously described. Among seven colonies analyzed, one containing the wild-type sequence was obtained, and designated pCA3D.

Table III

Amino Acid Composition of Wild-Type Cecropin-A Produced by pCA3D in E. coli		
Amino Acid	Experimental Values	Expected Values
Asp	2.2	2
Thr	1.0	1
Glu	4.1	4
Pro	1.1	1
Gly	4.4	4
Ala	5.2	5
Val	3.4	4
Ile	4.3	5
Leu	0.7	1
Phe	1.1	1
Lys	7.3	7
Arg	1.0	1
Trp	ND	1
Ser	0.3	0
Cys	0	0
Met	0	0
Tyr	0	0
His	0	0
N.D. = not determined.		

The nucleic acid and amino acid sequence are shown in Figure 7.

EXAMPLE 4

Assay of Bactericidal Activity

Agar plates (8 cm in diameter) were prepared with 6 ml of TYE medium containing about 10^7 viable cells of test organism. Wells, 2 mm diameter, were punched in the plates. The test material was dissolved in 50 mM phosphate buffer (pH 6.6) and 3 µl was applied to each well. The diameters of the inhibition zones, or halos, around the wells were measured after overnight incubation at 25°C. To determine a standard of halo formation, CA 1-33 protein was used. 1 µg and 3 µg of CA 1-33 in well applications caused halos with diameters of 8 mm and 11 mm respectively, on media infused with E. coli strain JF568. 10 µg and 30 µg of cyanogen bromide-digested pCA3A araB-CA fusion protein caused halos with diameters 7 mm and 18 mm respectively.

Cyanogen bromide-digested fusion protein fractionated by HPLC was also applied to this assay. Various peaks were tested, and one specific peak was shown to be bioactive. Other bacterial and yeast strains were also tested in this assay. Cyanogen bromide-digested and undigested fusion protein samples were applied in the amounts described above in the testing of E. coli strain JF568, and the resulting halo diameters were measured.

Results are shown for all cecropins obtained in Table IV.

Table IV

Bioactivity of CNBr-digested pCA3A, pCA3B, pCA3D fusion protein isolates (5µg) and ampicillin (750 ng) on various bacterial strains, as measured by halo activity.				
Bacterial Strain	Halo Activity in Millimeters			
	pCA3A¥	pCA3B¥	pCA3D*	Ampicillin
Staphylococcus aureus	NE	NE	NE	11.5
Streptococcus faecalis	NE	NE	NE	7.0
Salmonella derby	5.0	5.0	5.0	6.0
Pseudomonas PS-9	2.0	2.25	2.0	NE
Klebsiella pneumoniae	4.5	4.0	4.5	NE
Erwinia carotovora EC	5.0	4.5	4.75	6.5
Streptococcus pyogenes	NE	NE	NE	12.5
Xenohabdus nematophillus	3.5	3.0	2.75	12.5
Serratia marcescens	1.0	1.0	1.0	NE
E. coli 5506	5.0	4.5	5.0	4.5
E. coli 5506 DR	7.0	7.0	6.5	4.5

* Wild
¥ Mutants
NE No halo activity

EXAMPLE 5

Fermentative Production of Cecropin and of Tumor Growth Factor

This Example describes the fermentation production of alpha-TGF (tumor growth factor) and of cecropin under the control of araB promoter in E. coli.

The E. coli strain MC1061 contains plasmid-borne araB promoter regulating araB-alpha-TGF (pTGF58) or araB-cecropin (pCA3D) fusion gene. The cultures were grown at 37°C, 250 rpm in 100 ml of TYE culture medium containing ampicillin (0.1 gm/l). The cultures were grown until late exponential phase, approximately 200 Klett units; red filter. The cultures were then transferred to a four liter baffled shake flask containing 900 ml fresh TYE medium. Incubation continued for three more hours prior to inoculation into 9 liters of production medium. The production medium comprises the ingredients listed in Table V:

Table V

Base Medium		Additives	
Ingredients	Level (g/liter)	Ingredients	Level (g/liter)
casein hydrolysate	30	glycerol	16
yeast extract	1	$CaCl_2:2H_2O$	0.022
KH_2PO_4	3	$MgSO_4:7H_2O$	0.25
Na_2HPO_4	6	thiamine-HCl	0.01
NaCl	0.5	ampicillin	0.1
NH_4Cl	4		

Initially, the production fermentation conditions were set at 37°C, 800 rpm, and 1 vvm (volume of vessel per minute). As the dissolved oxygen was consumed during the growth time, both agitation and aeration rates were increased accordingly to maintain a minimum dissolved oxygen (D.O.) at 30%. The pH of this medium was consistently self-adjusted at 6.5-6.8, optimum for E. coli growth, such that acid and base control of the medium pH became unnecessary. Approximately four hours following inoculation into the production medium, cell density reach an $O.D._{600}$ of 10 when L-arabinose (50 g) was added to induce the synthesis of cecropin fusion protein. To further ensure the stability of the expression vector, 1 g of ampicillin was supplemented to the broth at an $O.D._{600}$ of 20.

Both alpha-TGF and cecropin-araB fusion proteins were localized only in the insoluble fraction of the E. coli cells. Microscopic examination showed insoluble inclusion bodies were formed inside the cells one hour following induction and were stably maintained throughout the fermentation. More than 95% of the cells contained at least one inclusion body which enlarged rapidly as the cell density kept increasing. The yield of cell mass was 13.1 g (dry weight) per liter. From the cell mass, approximately 30% was the fusion protein.

EXAMPLE 6

Construction Of E. coli Vector In Which Expression Of araB-hTGF Fusion Protein Is Under Regulation Of S. typhimurium araB Promoter

The araB promoter is positively regulated by the araC gene product. L-arabinose interacts with the araC protein to form an activator required for expression of the araB promoter. A plasmid which contains the S. typhimurium araB and araC genes was used to construct the araB-hTGF (human tumor growth factor alpha) expression vector. The strategy of plasmid construction is shown in Figure 8. The final plasmid phTGF5 contains an araB-hTGF fusion which codes for a protein of 548 amino acids. E. coli strain MC1061 containing phTGF5 was grown in minimal glycerol (1%) medium supplemented with casein hydrolysate (0.5%) and thiamine (1 µg/ml). When the density of the culture reached an A600 = 0.2, L-arabinose was added to 1% and incubation continued for five hours before the culture was harvested. A 55 KDal protein which represents approximately 10% of the total cellular protein was detected by SDS-PAGE.

EXAMPLE 7

Transcription Terminator 3' Added to the araB-hTGF gene

A transcription terminator is a DNA sequence which causes RNA polymerase to stop transcription. Placement of a transcription terminator at the 3' end of the araB-hTGF gene will prevent the expression of the undesired gene product(s) downstream from the transcription terminator. The strategy of plasmid construction as shown in Figure 9. The final plasmid phTGF58 contains an E. coli rrnB gene transcription terminator inserted at the 3'-end of the araB-hTGF gene. When the partially purified araB-hTGF proteins isolated from E. coli strain MC1061 containing either phTGF58 or phTGF5 (the parent plasmid) were compared after electrophoresis on SDS-polyacrylamide gel, a major contaminant protein, beta-lactamase, was significantly reduced in phTGF58-containing cells.

EXAMPLE 8

The araB Promoter and Calcitonin

A. The human calcitonin gene (hCT)

Calcitonin or thyrocalcitonin (CT) is the name given to the hypocalcemic hormone secreted from the thyroid. CT decreases bone resorptive activity. CT also acts on kidneys to stimulate increased urinary calcium and phosphate clearance. The rapid release of CT in response to blood calcium elevations suggests that the main purpose of CT is to protect higher animals from acute hypercalcemic episodes. Research in CT is examining its use in the control of Paget's disease, a problem of accelerated bone remodeling.

The sequence for the human gene for calcitonin is as follows (Craig et al., Nature, 295 345-347 (1982)):

```
Cys Gly Asn Leu Ser Thr Cys Met Leu Gly
TGC GGT AAT CTG AGT ACT TGC ATC CTG GGC

Thr Tyr Thr Gln Asp Phe Asn Lys Phe His
ACA TAC ACG CAG GAC TTC AAC AAG TTT CAC

Thr Phe Pro Gln Thr Ala Ile Gly Val Gly
ACG TTC CCC CAA ACT GCA ATT GGG GTT GGA

Ala Pro Gly
GCA CCT GGA
```

B. Construction of a plasmid carrying an inducible regulon, the araB structural gene and the hCT Gene

Plasmid p115, its construction shown in Figure 10, contains the correct calcitonin gene sequence. The plasmid was digested with endonucleases MstI and PstI to excise the calcitonin gene, as shown in Figure 11. The excised fragment was ligated to a plasmid carrying the araB structural gene (plasmid pING1, see Figure 12).

Plasmid pMH6 (Horwitz et al., Gene, 14: 309-319 (1981)) which contains intact araC and araB was used as starting material for construction of pING1. The construction scheme is shown in Figure 12. The plasmid pMH6 was digested with restriction endonuclease EcoRI and SalI, which cut in araB and araA gene, respectively, and the 1.9 kb SalI-EcoRI fragment was replaced by a 16 base-pair SalI-EcoRI fragment from M13 mp9 (Vieira and Messing, Gene, 19: 259-268 (1982)) to generate pING1. The reading frame of the araB gene was determined by fusion of the lacZ gene to the EcoRI site in pING1. Since both restriction endonuclease SmaI and MstI create blunt end fragments, the 760 base-pair fragment of pING1 was then replaced by a MstI-PstI fragment from plasmid #115 to generate phCT1. Plasmid #115 has the chemically synthesized hCT gene located in the MstI-PstI fragment. The joining of the MstI and SmaI end fused the hCT gene to the araB gene making an araB-hCT protein fusion.

C. Growth of E. coli cells containing phCT1

Transformed E. coli cells containing phCT1 were grown at 37°C, with shaking, to a density of 10^8-10^9/ml, and the synthesis of the araB-calcitonin fused protein was induced by the addition of L-arabinose (1% wt/vol) during additional growth for 1-6 hours. The harvested cells were sonicated (5 sec, 3 times) and the concentration of araB-calcitonin was assayed by an ELISA procedure using anti-hCT antibodies.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the invention, as limited by the scope of the appended claims.

Claims

Claims for the following Contracting States : BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. A replicable expression vehicle comprising a polynucleotide molecule expressible in a given host which comprises the sequence of the araB promoter operably linked to a gene which is heterologous to the host, wherein the gene codes for a polypeptide that is biologically active or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide, where the biologically active polypeptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not

thaumatin.

2. An expression vehicle as claimed in claim 1 which is a plasmid or bacteriophage.

3. An expression vehicle as claimed in claim 1 or 2 wherein the active polypeptide is human calcitonin, human tumour growth factor or cecropin.

4. An expression vehicle as claimed in any preceding claim wherein the cleavage site is Met, Asn-Gly, Asp-Pro, Lys, Arg or Lys-Arg.

5. A host transformed with an expression vehicle as claimed in any of claims 1 to 4.

6. A host as claimed in claim 5 which is *E. coil, S. typhimurium* or *Serratia marcescens.*

7. A process for producing a biologically active heterologous polypeptide in a bacterium transformed to express said peptide comprising:
(a) inserting a polynucleotide molecule comprising the araB promoter operably linked to a heterologous gene which codes for the biologically active polypeptide or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide, where the biologically active peptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not thaumatin, into a replicable expression vehicle;
(b) transforming a bacterium capable of expressing said polypeptide with said expression vehicle containing said polynucleotide molecule;
(c) culturing said transformed bacterium under culturing conditions; and
(d) recovering said polypeptide from said culture.

8. A process for producing a biologically active heterologous polypeptide by culturing a bacterium transformed with a polynucleotide molecule comprising the araB promoter operably linked to the gene heterologous to the bacterium and coding for the polypeptide, the process comprising the steps of:
(a) incubating a transformed bacterium in a culture medium under selected growth conditions until late exponential growth phase is reached, said bacterium being transformed with an expression vehicle as claimed in any of claims 1 to 4;
(b) transferring the culture of step (a) to fresh culture medium and incubating the culture for a period of time sufficient to allow further growth;
(c) inoculating the culture of step (b) with production medium and incubating the culture at fermentation conditions to achieve maximal bacterial growth and for a time sufficient to achieve a selected cell density;
(d) adding an amount of L-arabinose effective to produce substantial levels of the heterologous polypeptide in the medium; and
(e) recovering the heterologous polypeptide.

Claims for the following Contracting State : AT

1. A process for producing a biologically active heterologous polypeptide in a bacterium transformed to express said polypeptide comprising:
(a) inserting a polynucleotide molecule comprising the araB promoter operably linked to a heterologous gene, which codes for the biologically active peptide or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide into a replicable expression vehicle where the biologically active polypeptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not thaumatin;
(b) transforming a bacterium capable of expressing said polypeptide with said expression vehicle containing said polynucleotide molecule;
(c) culturing said transformed bacterium under culturing conditions; and
(d) recovering said polypeptide from said culture.

2. A process as claimed in claim 1 wherein the expression vehicle is a plasmid or bacteriophage.

3. A process for producing a biologically active heterologous polypeptide by culturing a bacterium transformed with a polynucleotide molecule comprising the araB promoter operably linked to a gene heterologous to the bacterium and coding for the polypeptide, the process comprising the steps of:
(a) incubating a transformed bacterium in a culture medium under selected growth conditions until late exponential growth phase is reached, said bacterium being transformed with a polynucleotide molecule comprising the araB promoter operably linked to the heterologous gene and coding for a biologically active polypeptide or a fusion precursor protein comprising a second polypeptide and a selective cleavage site adjacent the biologically active polypeptide the biologically active polypeptide into a replicable expression vehicle where the biologically active polypeptide is an enzyme, hormone, antibody, haemoglobin, structural protein, α, β or γ interferon, interleukin, insulin or tissue plasminogen activator, but not thaumatin;
(b) transferring the culture of step (a) to fresh culture medium and incubating said culture for a period of time sufficient to allow further growth;
(c) inoculating said culture of step (b) with production medium and incubating said culture at fermentation conditions to achieve maximal bacterial growth and for a time sufficient to achieve a selected cell density;
(d) adding an amount of L-arabinose effective to produce substantial levels of said heterologous polypeptide in said medium; and
(e) recovering said heterologous polypeptide.

4. A process as claimed in any preceding claim wherein the active polypeptide is human calcitonin, human tumour growth factor or cecropin.

5. A process as claimed in any preceding claim wherein the cleavage site is Met, Asn-Gly, Asp-Pro, Lys, Arg or Lys-Arg.

6. A process as claimed in any preceding claim wherein the bacterium is *E. coli, S. typhimurium* or *Serratia marcescens.*

Patentansprüche
Patentansprüche für folgende Vertragsstaaten : BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. Replizierbarer Expressionsvektor, bestehend aus einem Polynucleotidmolekül, das in einem bestimmten Wirt exprimierbar ist und die Sequenz des araB-Promotors umfaßt, der mit einem Gen funktionsfähig verbunden ist, das zu dem Wirt heterolog ist, wobei das Gen für ein Polypeptid kodiert, das biologisch aktiv ist, oder für ein Fusionsvorläuferprotein, das ein zweites Polypeptid und eine selektive Spaltstelle angrenzend an das biologisch aktive Polypeptid umfaßt, wobei das biologisch aktive Polypeptid ein Enzym, Hormon, Antikörper, Hämoglobin, Strukturprotein, α-, β- oder γ-Interferon, Interleukin, Insulin oder Gewebsplasminogenaktivator, aber nicht Thaumatin ist.

2. Expressionsvektor nach Anspruch 1, der ein Plasmid oder Bakteriophage ist.

3. Expressionsvektor nach Anspruch 1 oder 2, wobei das aktive Polypeptid menschliches Calcitonin, menschlicher Tumorwachstumsfaktor oder Cecropin ist.

4. Expressionsvektor nach einem der vorangehenden Ansprüche, wobei die Spaltstelle Met, Asn-Gly, Asp-Pro, Lys, Arg oder Lys-Arg ist.

5. Wirt, der mit einem Expressionsvektor nach einem der Ansprüche 1 bis 4 transformiert wurde.

6. Wirt nach Anspruch 5, der E. coli, S. typhimurium oder Serratia marcescens ist.

7. Verfahren zur Herstellung eines biologisch aktiven, heterologen Polypeptids in einem zur Expression des Peptids transformierten Bakterium, bestehend aus:
a) dem Einsetzen eines Polynucleotidmoleküls, bestehend aus dem araB-Promotor, der mit einem heterologen Gen funktionsfähig verbunden ist, das für das biologisch aktive Polypeptid oder ein Fusionsvorläuferprotein codiert, das ein zweites Polypeptid und eine selektive Spaltstelle umfaßt, die an das biologisch aktive Polypeptid angrenzt, in einen replizierbaren Expressionsvektor, wobei das

biologisch aktive Polypeptid ein Enzym, Hormon, Antikörper, Hämoglobin, Strukturprotein, α-, β- oder γ-Interferon, Interleukin, Insulin oder Gewebsplasminogenaktivator, aber nicht Thaumatin ist;
b) dem Transformieren eines Bakteriums, das zur Expression des Polypeptids fähig ist, mit dem Expressionsvektor, der das Polynucleotidmolekül enthält;
c) dem Kultivieren des transformierten Bakteriums unter Kultivierungsbedingungen; und
d) dem Gewinnen des Polypeptids aus der Kultur.

8. Verfahren zur Herstellung eines biologisch aktiven, heterologen Polypeptids durch Kultivieren eines Bakteriums, das mit einem Polynucleotidmolekül transformiert ist, das den araB-Promotor umfaßt, der mit dem Gen funktionsfähig verbunden ist, das zu dem Bakterium heterolog ist, und für das Polypeptid codiert, wobei das Verfahren folgende Schritte umfaßt:
a) das Inkubieren eines transformierten Bakteriums in einem Kulturmedium unter bestimmten Wachstumsbedingungen, bis die späte exponentielle Wachstumsphase erreicht ist, wobei das Bakterium mit einem Expressionsvektor nach einem der Ansprüche 1 bis 4 transformiert ist;
b) dem Einbringen der Kultur von Schritt a) in ein frisches Kulturmedium und dem Inkubieren der Kultur über einen ausreichenden Zeitraum, der ein weiteres Wachstum ermöglicht;
c) dem Beimpfen der Kultur von Schritt b) mit Produktionsmedium und dem Inkubieren der Kultur unter Fermentationsbedingungen zum Erzielen eines maximalen bakteriellen Wachstums über einen ausreichenden Zeitraum, der zur Erzielung einer bestimmten Zelldichte ausreicht;
d) dein Zugeben einer Menge von L-Arabinose, welche die Herstellung wesentlicher Mengen des heterologen Polypeptids in dem Medium bewirkt; und
e) dem Gewinnen des heterologen Peptids.

Patentansprüche für folgenden Vertragsstaat : AT

1. Verfahren zur Herstellung eines biologisch aktiven, heterologen Polypeptids in einem zur Expression des Polypeptids transformierten Bakterium, bestehend aus:
a) dem Einsetzen eines Polynucleotidmoleküls, bestehend aus dem araB-Promotor, der mit einem heterologen Gen funktionsfähig verbunden ist, das für das biologisch aktive Polypeptid oder ein Fusionsvorläuferprotein codiert, das ein zweites Polypeptid und eine selektive Spaltstelle umfaßt, die an das biologisch aktive Polypeptid angrenzt, in einen replizierbaren Expressionsvektor, wobei das biologisch aktive Polypeptid ein Enzym, Hormon, Antikörper, Hämoglobin, Strukturprotein, α-, β- oder γ-Interferon, Interleukin, Insulin oder Gewebsplasminogenaktivator, aber nicht Thaumatin ist;
b) dem Transformieren eines Bakteriums, das zur Expression des Polypeptids mit dem Expressionsvektor, der das Polynucleotidmolekül enthält, fähig ist;
c) dem Kultivieren des transformierten Bakteriums unter Kultivierungsbedingungen; und
d) dem Gewinnen des Polypeptids aus der Kultur.

2. Verfahren nach Anspruch 1, wobei der Expressionsvektor ein Plasmid oder Bakteriophage ist.

3. Verfahren zur Herstellung eines biologisch aktiven, heterologen Polypeptids durch Kultivieren eines Bakteriums, das mit einem Polynucleotidmolekül transformiert ist, das den araB-Promotor umfaßt, der mit einem Gen funktionsfähig verbunden ist, das zu dem Bakterium heterolog ist und für das Polypeptid codiert, wobei das Verfahren folgende Schritte umfaßt:
a) das Inkubieren eines transformierten Bakteriums in einem Kulturmedium unter bestimmten Wachstumsbedingungen bis die späte exponentielle Wachstumsphase erreicht ist, wobei das Bakterium mit einem Polynucleotidmolekül transformiert ist, das den araB-Promotor enthält, der mit dem heterologen Gen funktionsfähig verbunden ist und für ein biologisch aktives Polypeptid oder ein Fusionsvorläuferprotein codiert, das ein zweites Polypeptid und eine selektive Spaltstelle umfaßt, die an das biologisch aktive Polypeptid angrenzt, in einen replizierbaren Expressionsvektor, wobei das biologisch aktive Polypeptid ein Enzym, Hormon, Antikörper, Hämoglobin, Strukturprotein, α-, β- oder γ-Interferon, Interleukin, Insulin oder Gewebsplasminogenaktivator, aber nicht Thaumatin ist;
b) dem Einbringen der Kultur von Schritt a) in ein frisches Kulturmedium und dem Inkubieren der Kultur über einen ausreichenden Zeitraum, der ein weiteres Wachstum ermöglicht;
c) dem Beimpfen der Kultur von Schritt b) mit Produktionsmedium und dem Inkubieren der Kultur unter Fermentationsbedingungen zur Erzielung eines maximalen bakteriellen Wachstums über einen ausreichenden Zeitraum, der zur Erzielung einer bestimmten Zelldichte ausreicht;

d) dem Zugeben einer Menge von L-Arabinose, welche die Herstellung wesentlicher Mengen des heterologen Polypeptids in dem Medium bewirkt; und
e) dem Gewinnen des heterologen Peptids.

4. Verfahren nach einem der vorangehenden Ansprüche, wobei das aktive Polypeptid menschliches Calcitonin, menschlicher Tumorwachstumsfaktor oder Cecropin ist.

5. Verfahren nach einem der vorangehenden Ansprüche, wobei die Spaltstelle Met, Asn-Gly, Asp-Pro, Lys, Arg oder Lys-Arg ist.

6. Verfahren nach einem der vorangehenden Ansprüche, wobei das Bakterium E. coli, S. typhimurium oder Serratia marcescens ist.

Revendications

Revendications pour les Etats contractants suivants : BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. Véhicule d'expression réplicable comprenant une molécule de polynucléotide exprimable dans un hôte donné qui comprend la séquence du promoteur d'araB ligaturé de manière opérable à un gène qui est héterologue à l'hôte, où le gène code pour un polypeptide qui est biologiquement actif ou une protéine de fusion précurseur comprenant un second polypeptide et un site de coupure sélectif adjacent au peptide biologiquement actif, où le peptide biologiquement actif est une enzyme, une hormone, un anticorps, une hémoglobine, une protéine de structure, un interféron α, β ou γ, une interleukine, une insuline ou un activateur de plasminogène tissulaire mais pas la thaumatine.

2. Véhicule d'expression suivant la revendication 1, qui est un plasmide ou un bactériophage.

3. Véhicule d'expression suivant la revendication 1 ou 2, où le polypeptide actif est la calcitonine humaine, le facteur de croissance tumorale humaine ou la cécropine.

4. Véhicule d'expression suivant l'une quelconque des revendications précédentes, où le site de coupure est Met, Asn-Gly, Asp-Pro, Lys, Arg ou Lys-Arg.

5. Hôte transformé avec un véhicule d'expression suivant l'une quelconque des revendications 1 à 4.

6. Hôte suivant la revendication 5, qui est Escherichia coli, Salmonella typhimurium ou Serratia marcescens.

7. Procédé de production d'un polypeptide hétérologue biologiquement actif dans une bactérie transformée pour exprimer ledit peptide comprenant :
(a) l'insertion d'une molécule de polynucléotide comprenant le promoteur d'araB ligaturé de manière opérable à un gène hétérologue qui code pour le polypeptide biologiquement actif ou une protéine de fusion précurseur comprenant un second polypeptide et un site de coupure sélectif adjacent au polypeptide biologiquement actif, où le polypeptide biologiquement actif est une enzyme, une hormone, un anticorps, une hémoglobine, une protéine de structure, un interféron α, β ou γ, une interleukine, une insuline ou un activateur de plasminogène tissulaire, mais pas la thaumatine, dans un véhicule d'expression réplicable;
(b) la transformation d'une bactérie capable d'exprimer ledit polypeptide avec ledit véhicule d'expression contenant ladite molécule de polynucléotide;
(c) la culture de ladite bactérie transformée sous des conditions de culture, et
(d) la récupération dudit polypeptide à partir de la ladite culture.

8. Procédé de production d'un polypeptide hétérologue biologiquement actif par culture d'une bactérie transformée avec une molécule de polynucléotide comprenant le promoteur d'araB ligaturé de manière opérable au gène hétérologue à la bactérie et codant pour le polypeptide, le procédé comprenant les étapes de :
(a) incubation d'une bactérie transformée dans un milieu de culture sous des conditions de croissance choisies jusqu'à ce que la phase exponentielle tardive de croissance soit atteinte, ladite bactérie étant transformée avec un véhicule d'expression suivant l'une quelconque des revendica-

tions 1 à 4;
(b) transfert de la culture de l'étape (a) dans un milieu de culture frais et incubation de la culture pendant une période de temps suffisante pour permettre une croissance supplémentaire;
(c) inoculation de la culture de l'étape (b) avec un milieu de production et incubation de la culture à des conditions de fermentation pour atteindre la croissance bactérienne maximale et pendant un temps suffisant pour atteindre une densité cellulaire choisie;
(d) addition d'une quantité de L-arabinose efficace pour produire des taux substantiels de polypeptide hétérologue dans le milieu, et
(e) récupération du polypeptide hétérologue.

Revendications pour l'Etat contractant suivant : AT

1. Procédé de production d'un polypeptide hétérologue biologiquement actif dans une bactérie transformée pour exprimer ledit polypeptide comprenant :
(a) l'insertion d'une molécule de polynucléotide comprenant le promoteur d'araB ligaturé de manière opérable à un gène hétérologue qui code pour le peptide biologiquement actif ou une protéine de fusion précurseur comprenant un second polypeptide et un site de coupure sélectif adjacent au polypeptide biologiquement actif, dans un véhicule d'expression réplicable où le polypeptide biologiquement actif est une enzyme, une hormone, un anticorps, une hémoglobine, une protéine de structure, un interféron α, β ou γ, une interleukine, une insuline ou un activateur de plasminogène tissulaire, mais pas la thaumatine;
(b) la transformation d'une bactérie capable d'exprimer ledit polypeptide avec ledit véhicule d'expression contenant ladite molécule du polynucléotide;
(c) la culture de ladite bactérie transformée dans des conditions de culture, et
(d) la récupération dudit polypeptide à partir de la ladite culture.

2. Procédé suivant la revendication 1, où le véhicule d'expression est un plasmide ou un bactériophage.

3. Procédé de production d'un polypeptide hétérologue biologiquement actif par culture d'une bactérie transformée avec une molécule de polynucléotide comprenant le promoteur d'araB ligaturé de manière opérable au gène hétérologue à la bactérie et codant pour le polypeptide, le procédé comprenant les étapes de :
(a) incubation d'une bactérie transformée dans un milieu de culture sous des conditions de croissance choisies jusqu'à ce que la phase exponentielle ultime de croissance soit atteinte, ladite bactérie étant transformée avec une molécule de polynucléotide comprenant le promoteur d'araB ligaturé de manière opérable à un gène hétérologue qui code pour un peptide biologiquement actif ou une protéine de fusion précurseur comprenant un second polypeptide et un site de coupure sélectif adjacent au polypeptide biologiquement actif, dans un véhicule d'expression réplicable où le polypeptide biologiquement actif est une enzyme, une hormone, un anticorps, une hémoglobine, une protéine de structure, un interféron α, β ou γ, une interleukine, une insuline ou un activateur de plasminogène tissulaire mais pas la thaumatine;
(b) transfert de la culture de l'étape (a) dans un milieu de culture frais et incubation de la culture pendant une période de temps suffisante pour permettre une croissance supplémentaire;
(c) inoculation de la culture de l'étape (b) avec un milieu de production et incubation de ladite culture à des conditions de fermentation pour atteindre la croissance bactérienne maximale et pour un temps suffisant pour atteindre la densité cellulaire choisie;
(d) addition d'une quantité de L-arabinose efficace pour produire des taux substantiels dudit polypeptide hétérologue dans ledit milieu, et
(e) récupération du polypeptide hétérologue.

4. Procédé suivant l'une quelconque des revendications précédentes, où le polypeptide actif est la calcitonine humaine, le facteur de croissance tumorale humaine ou la cécropine.

5. Procédé suivant l'une quelconque des revendications précédentes, où le site de coupure est Met, Asn-Gly, Asp-Pro, Lys, Arg ou Lys-Arg.

6. Procédé suivant l'une quelconque des revendications précédentes, où la bactérie est Escherichia coli, Salmonella typhimurium ou Serratia marcescens.

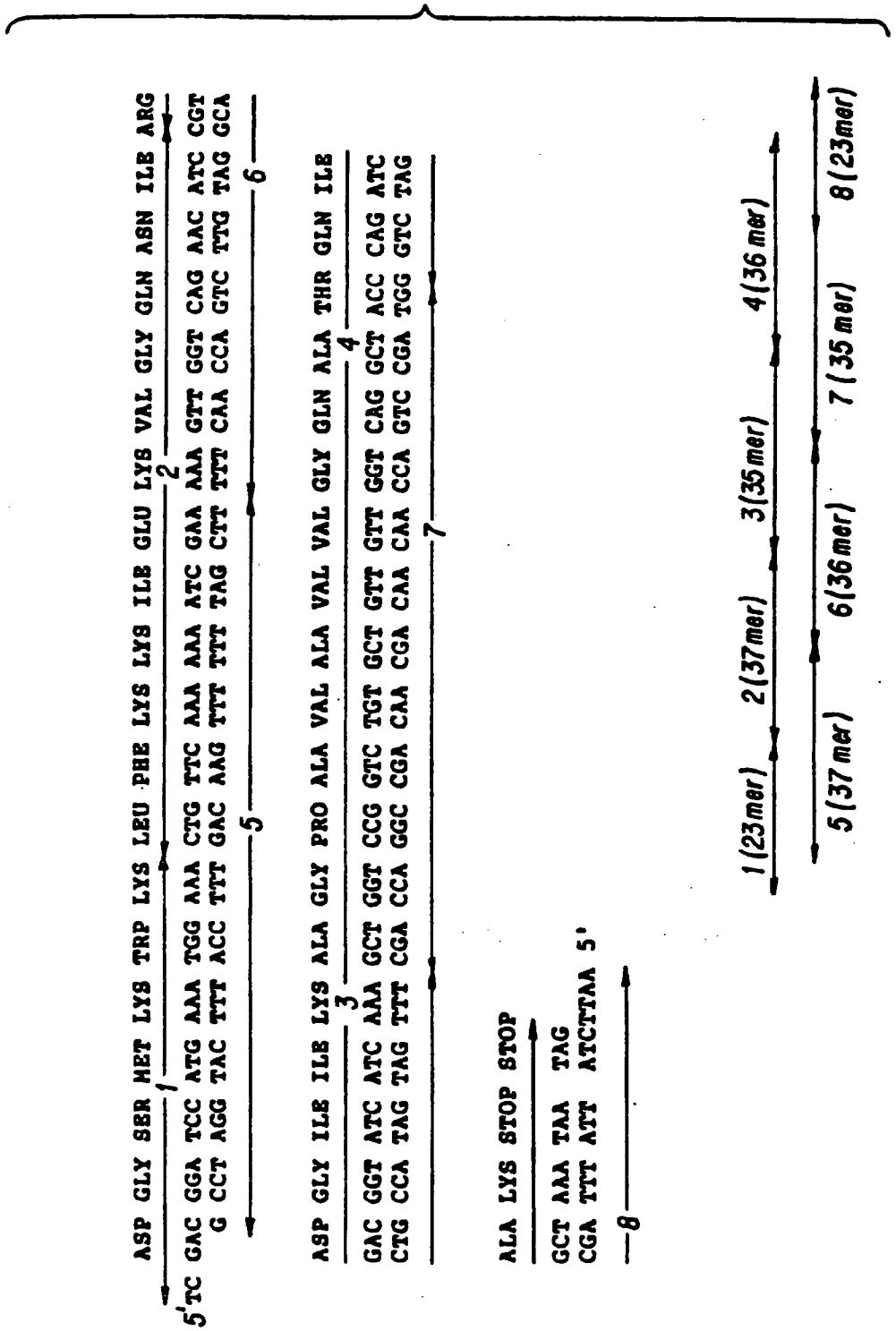
ASP GLY SER MET LYS TRP LYS LEU PHE LYS LYS ILE GLU LYS VAL GLY GLN ASN ILE ARG
1
2
5'TC GAC GGA TCC ATG AAA TGG AAA CTG TTC AAA AAA ATC GAA AAA GTT GGT CAG AAC ATC CGT
G CCT AGG TAC TTT ACC TTT GAC AAG TTT TTT TAG CTT TTT CAA CCA GTC TTG TAG GCA
5
6
ASP GLY ILE ILE LYS ALA GLY PRO ALA VAL ALA VAL VAL GLY GLN ALA THR GLN ILE
3
4
GAC GGT ATC ATC AAA GCT GGT CCG GTC TGT GCT GTT GTT GGT CAG GCT ACC CAG ATC
CTG CCA TAG TAG TTT CGA CCA GGC CGA CAA CGA CAA CAA CCA GTC CGA TGG GTC TAG
7
ALA LYS STOP STOP
GCT AAA TAA TAG
CGA TTT ATT ATCTTAA 5'
8
1(23mer)
2(37mer)
3(35mer)
4(36 mer)
5(37 mer)
6(36mer)
7(35 mer)
8(23mer)

FIG. 1

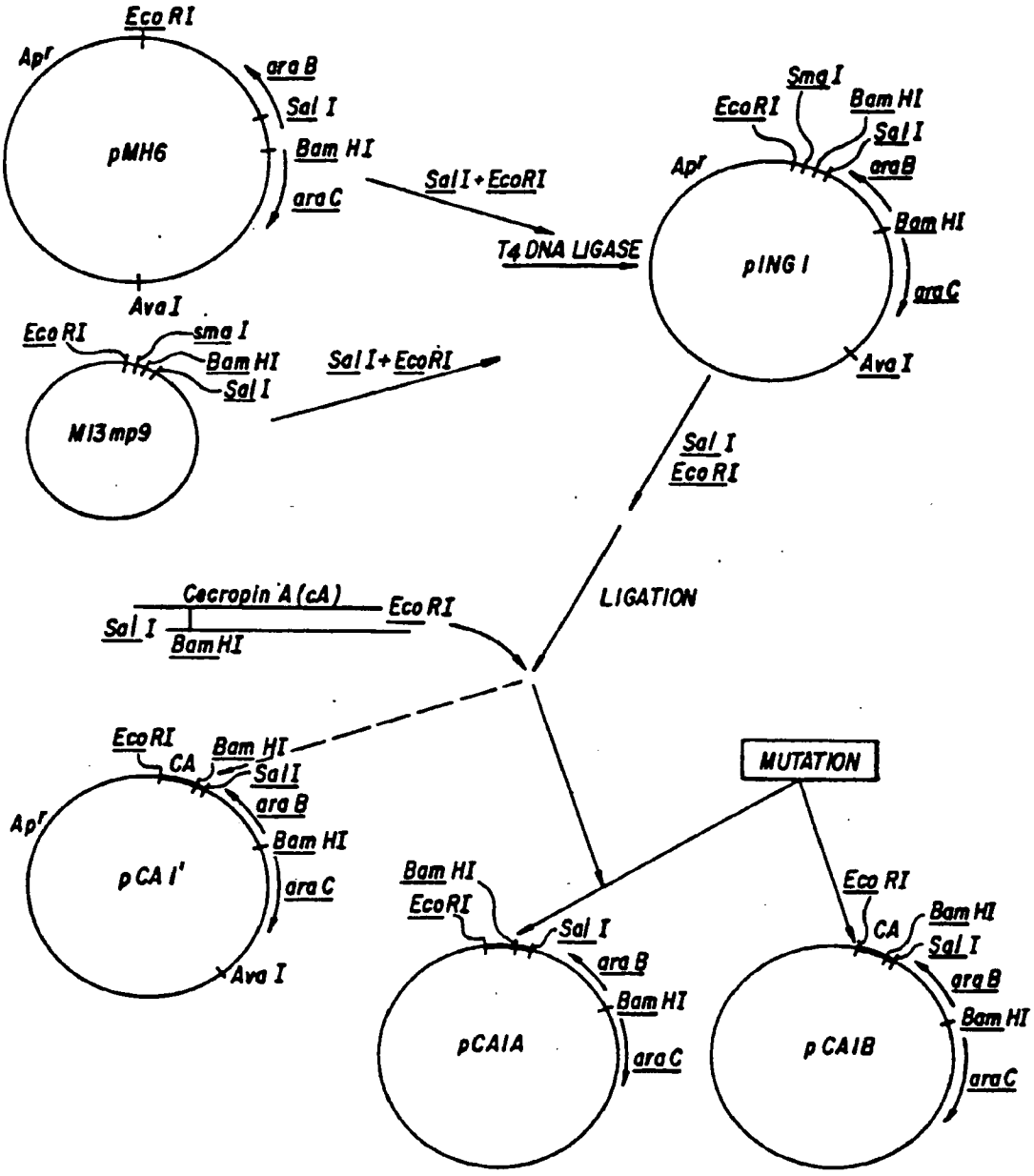
Eco RI
Ap^r
ara B
Sal I
pMH6
Bam HI
ara C
Ava I
Sal I + Eco RI
T4 DNA LIGASE
Smq I
Eco RI
Bam HI
Sal I
Ap^r
ara B
pING I
Bam HI
ara C
Ava I
Eco RI
sma I
Bam HI
Sal I
M13mp9
Sal I + Eco RI
Sal I
Eco RI
LIGATION
Cecropin A (cA)
Eco RI
Sal I
Bam HI
MUTATION
Eco RI
CA
Bam HI
Sal I
Ap^r
ara B
pCA I'
Bam HI
ara C
Ava I
Bam HI
Eco RI
Sal I
ara B
pCAIA
Bam HI
ara C
Eco RI
CA
Bam HI
Sal I
ara B
pCAIB
Bam HI
ara C

FIG. 2

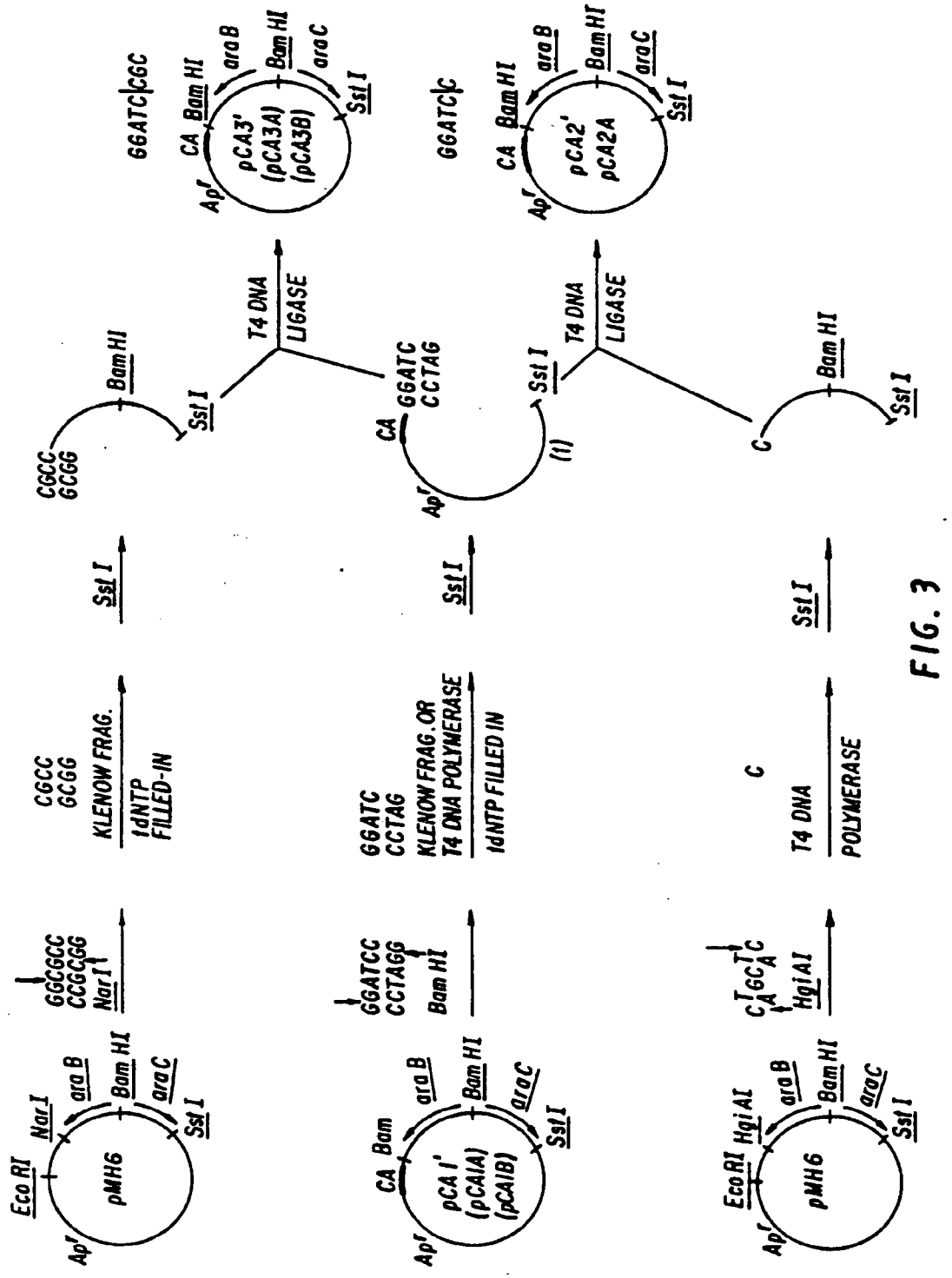
Eco RI
Nar I
ara B
Bam HI
ara C
Sst I
pMH6
Ap^r
GGCGCC
CCGCGG
Nar I
KLENOW FRAG.
IdNTP
FILLED-IN
CGCC
GCGG
Sst I
Bam HI
CGCC
GCGG
Sst I
T4 DNA
LIGASE
GGATC|CGC
CA Bam HI
ara B
Bam HI
ara C
Sst I
pCA3'
(pCA3A)
(pCA3B)
Ap^r
GGATCC
CCTAGG
Bam HI
CA Bam
ara B
Bam HI
ara C
Sst I
pCA 1'
(pCA1A)
(pCA1B)
Ap^r
GGATC
CCTAG
KLENOW FRAG. OR
T4 DNA POLYMERASE
IdNTP FILLED IN
Sst I
CA
GGATC
CCTAG
Sst I
(1)
Ap^r
T4 DNA
LIGASE
GGATC|C
CA Bam HI
ara B
Bam HI
ara C
Sst I
pCA2'
pCA2A
Ap^r
C
T4 DNA
POLYMERASE
Sst I
C
Bam HI
Sst I
C|GCA|C
Hgi AI
Eco RI
Hgi AI
ara B
Bam HI
ara C
Sst I
pMH6
Ap^r

FIG. 3

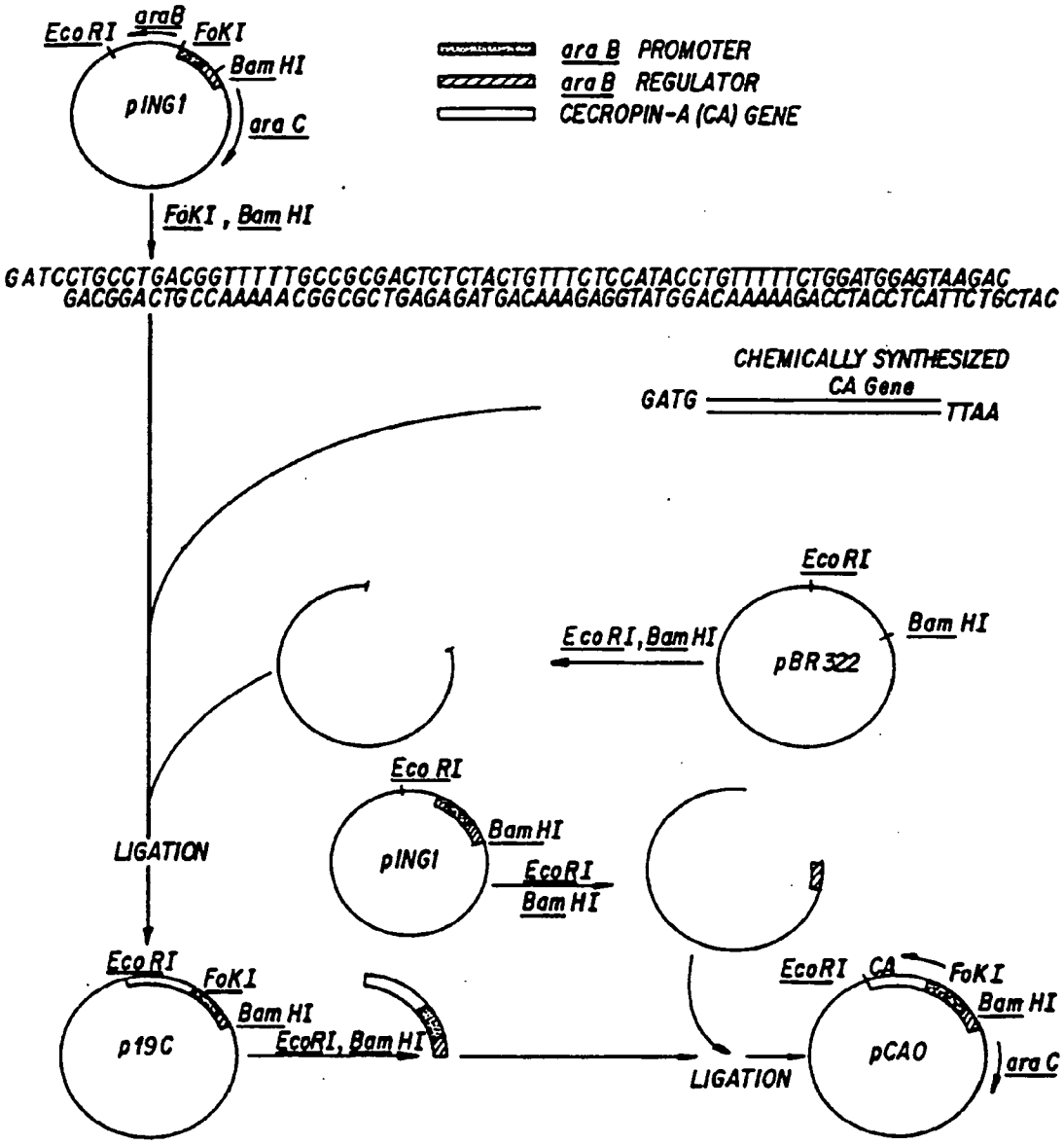
EcoRI
araB
FoKI
Bam HI
pING1
ara C
ara B PROMOTER
ara B REGULATOR
CECROPIN-A (CA) GENE
FoKI , Bam HI
GATCCTGCCTGACGGTTTTTGCCGCGACTCTCTACTGTTTCTCCATACCTGTTTTTCTGGATGGAGTAAGAC
GACGGACTGCCAAAAACGGCGCTGAGAGATGACAAAGAGGTATGGACAAAAAGACCTACCTCATTCTGCTAC
CHEMICALLY SYNTHESIZED
CA Gene
GATG
TTAA
EcoRI
Bam HI
EcoRI, Bam HI
pBR322
Eco RI
Bam HI
pING1
EcoRI
Bam HI
LIGATION
EcoRI
FoKI
Bam HI
p19C
EcoRI, Bam HI
EcoRI
CA
FoKI
Bam HI
pCAO
ara C
LIGATION

FIG. 4

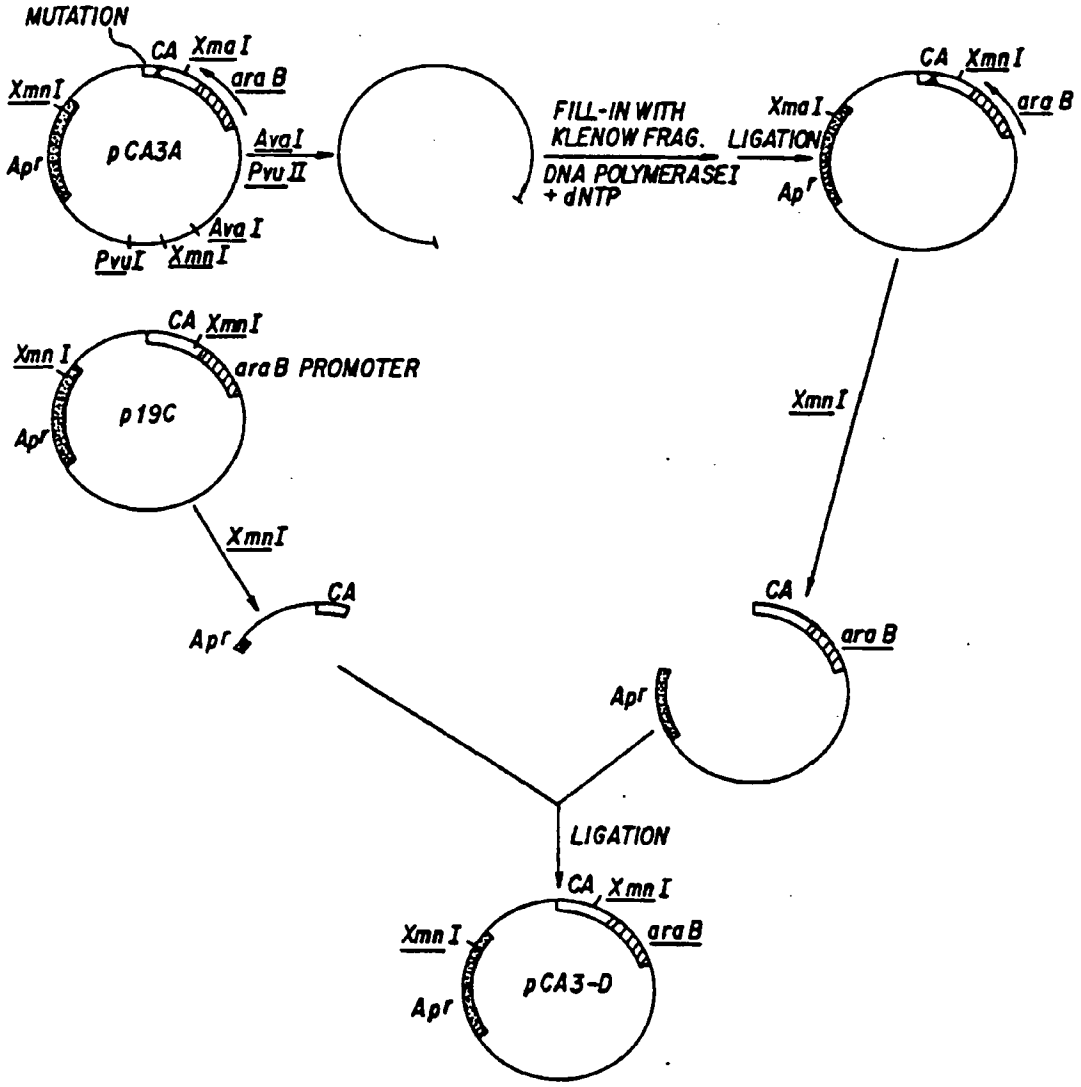
MUTATION
CA
Xma I
ara B
Xmn I
pCA3A
Apr
Ava I
Pvu II
Ava I
Pvu I
Xmn I
FILL-IN WITH
KLENOW FRAG.
DNA POLYMERASE I
+ dNTP
LIGATION
Xma I
CA
Xmn I
ara B
Apr
Xmn I
CA
Xmn I
ara B PROMOTER
p19C
Apr
Xmn I
CA
Apr
CA
ara B
Apr
LIGATION
CA
Xmn I
ara B
Xmn I
pCA3-D
Apr

FIG. 5

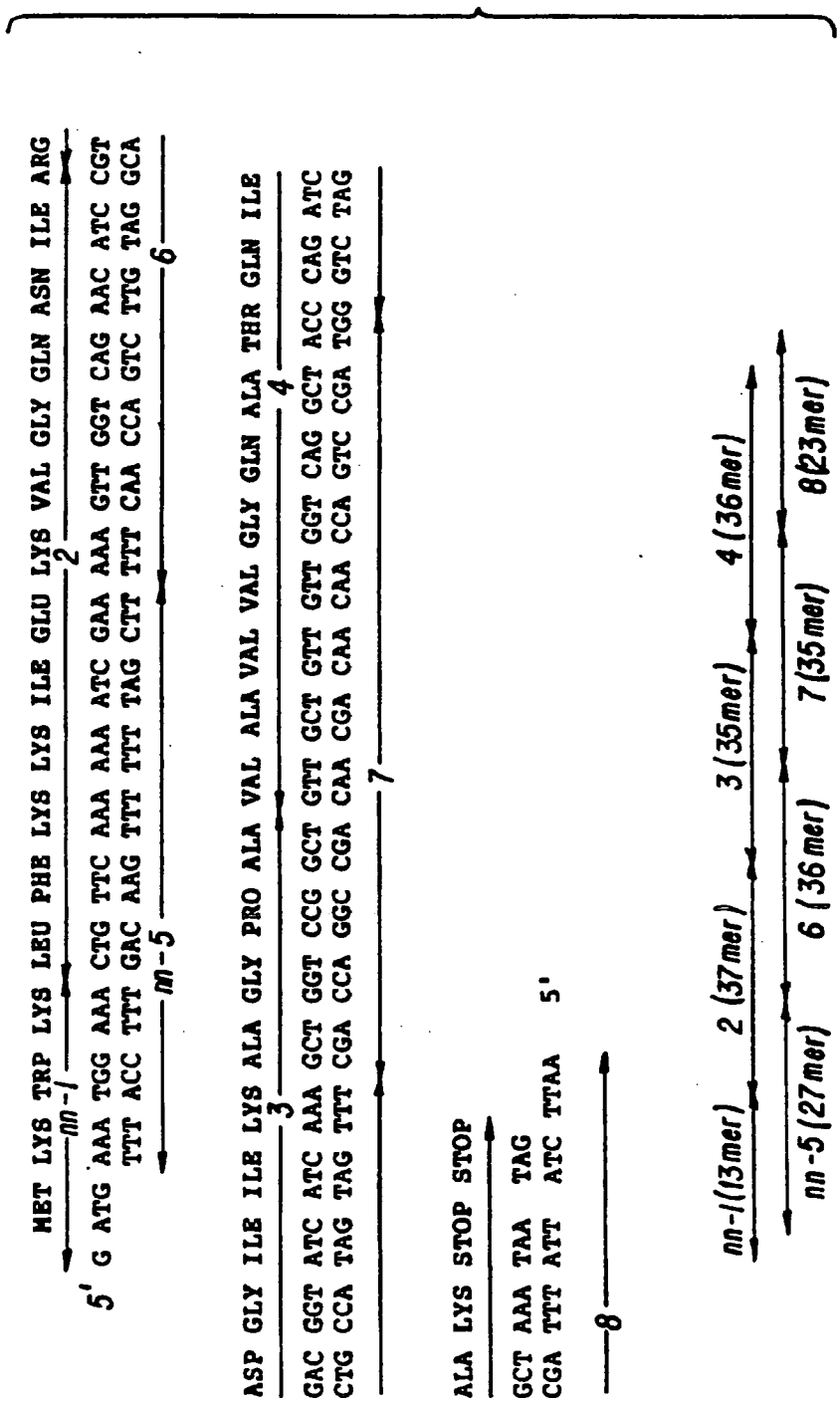
MET LYS TRP LYS LEU PHE LYS LYS ILE GLU LYS VAL GLY GLN ASN ILE ARG
nn-1
2
5' G ATG AAA TGG AAA CTG TTC AAA AAA ATC GAA AAA GTT GGT CAG AAC ATC CGT
TTT ACC TTT GAC AAG TTT TTT TAG CTT TTT CAA CCA GTC TTG TAG GCA
m-5
6
ASP GLY ILE ILE LYS ALA GLY PRO ALA VAL ALA VAL VAL GLY GLN ALA THR GLN ILE
3
4
GAC GGT ATC ATC AAA GCT GGT CCG GCT GTT GCT GTT GTT GGT CAG GCT ACC CAG ATC
CTG CCA TAG TAG TTT CGA CCA GGC CGA CAA CGA CAA CAA CCA GTC CGA TGG GTC TAG
7
ALA LYS STOP STOP
GCT AAA TAA TAG
CGA TTT ATT ATC TTAA 5'
8
nn-1(13mer)
2 (37mer)
3 (35mer)
4 (36mer)
nn-5 (27mer)
6 (36mer)
7 (35mer)
8 (23mer)

FIG. 6

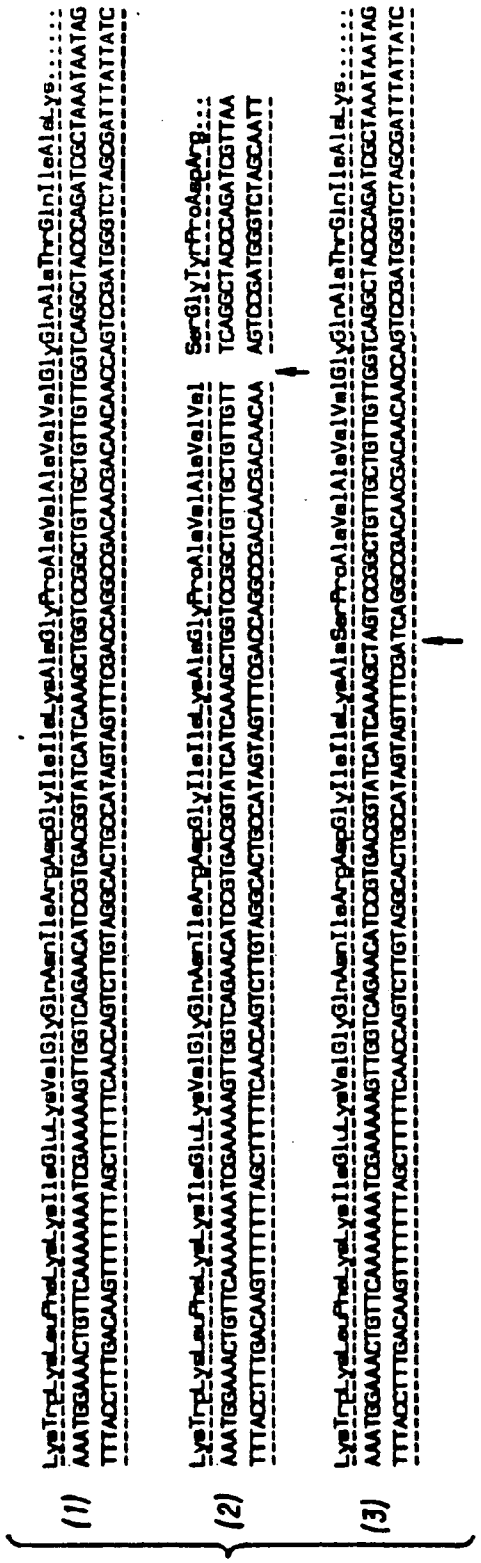

FIG.7

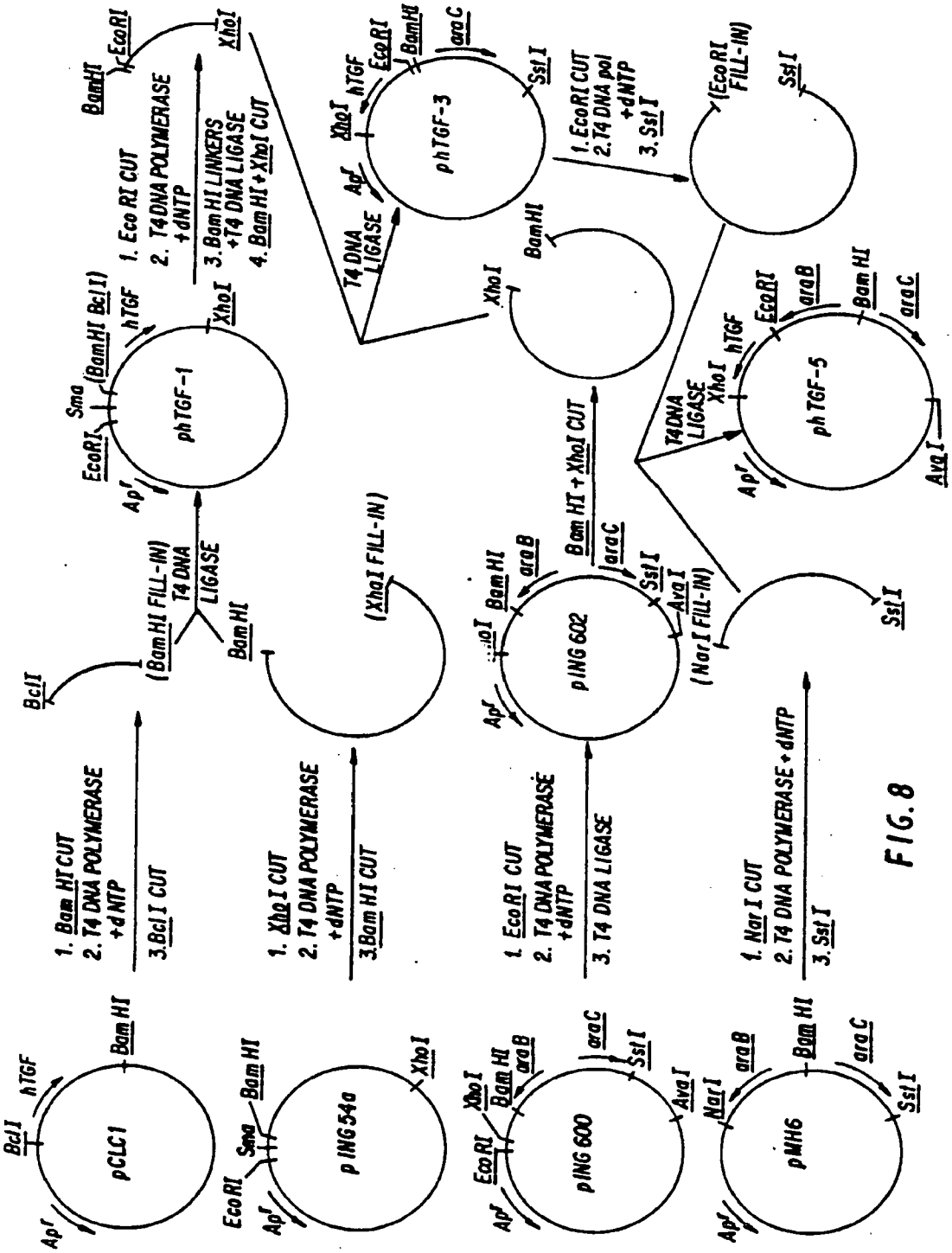
pCLC1
hTGF
Bam HI
Bcl I
Ap^r
1. Bam HI CUT
2. T4 DNA POLYMERASE +dNTP
3. Bcl I CUT
pING54a
Bam HI
Xho I
Eco RI
Sma
Ap^r
1. Xho I CUT
2. T4 DNA POLYMERASE +dNTP
3. Bam HI CUT
(Xho I FILL-IN)
pING600
Xho I
Bam HI
ara B
ara C
Sst I
Ava I
Eco RI
Ap^r
1. Eco RI CUT
2. T4 DNA POLYMERASE +dNTP
3. T4 DNA LIGASE
pMH6
Nar I
ara B
Bam HI
ara C
Sst I
Ap^r
1. Nar I CUT
2. T4 DNA POLYMERASE + dNTP
3. Sst I
(Nar I FILL-IN)
pING602
Bam HI + Xho I CUT
phTGF-1
Sma
EcoRI
(BamHI Bcl I)
hTGF
Xho I
Ap^r
1. Eco RI CUT
2. T4 DNA POLYMERASE +dNTP
3. Bam HI LINKERS +T4 DNA LIGASE
4. Bam HI + Xho I CUT
(Bam HI FILL-IN)
T4 DNA LIGASE
Bam HI
T4 DNA LIGASE
phTGF-3
Xho I
hTGF
Eco RI
Bam HI
ara C
Sst I
Ap^r
1. Eco RI CUT
2. T4 DNA pol +dNTP
3. Sst I
T4DNA LIGASE
phTGF-5
Xho I
hTGF
Eco RI
ara B
Bam HI
ara C
Ava I
Ap^r
(Eco RI FILL-IN)
Sst I

FIG. 8

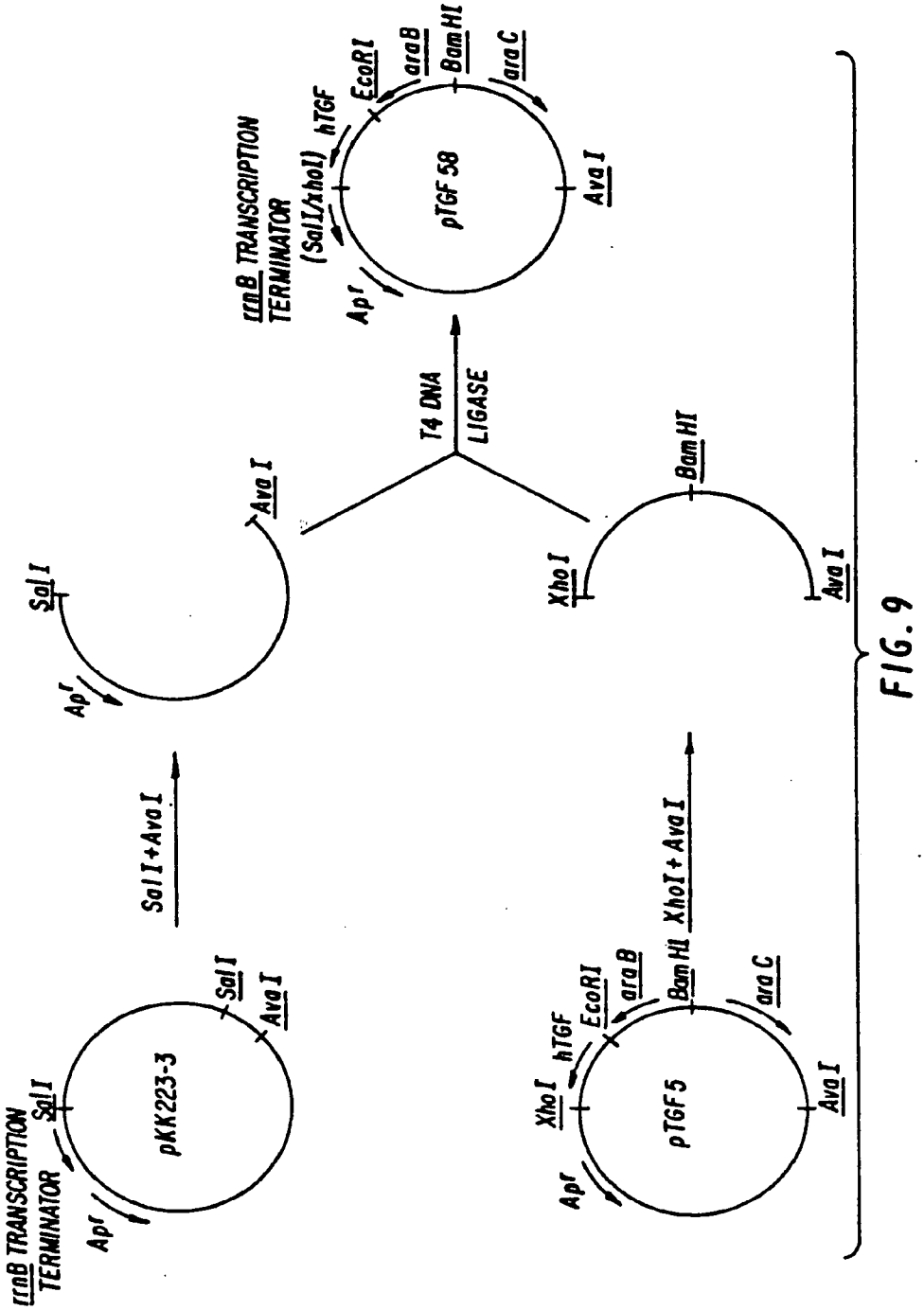
rrnB TRANSCRIPTION TERMINATOR
Sal I
Ap^r
pKK223-3
Sal I
Ava I
Sal I + Ava I
Sal I
Ap^r
Ava I
T4 DNA LIGASE
rrnB TRANSCRIPTION TERMINATOR
(Sal I/Xho I)
hTGF
EcoR I
ara B
Bam HI
ara C
Ava I
pTGF 58
Ap^r
Xho I
Bam HI
Ava I
Xho I
hTGF
EcoR I
ara B
Bam HI
Xho I + Ava I
ara C
Ava I
pTGF 5
Ap^r

FIG. 9

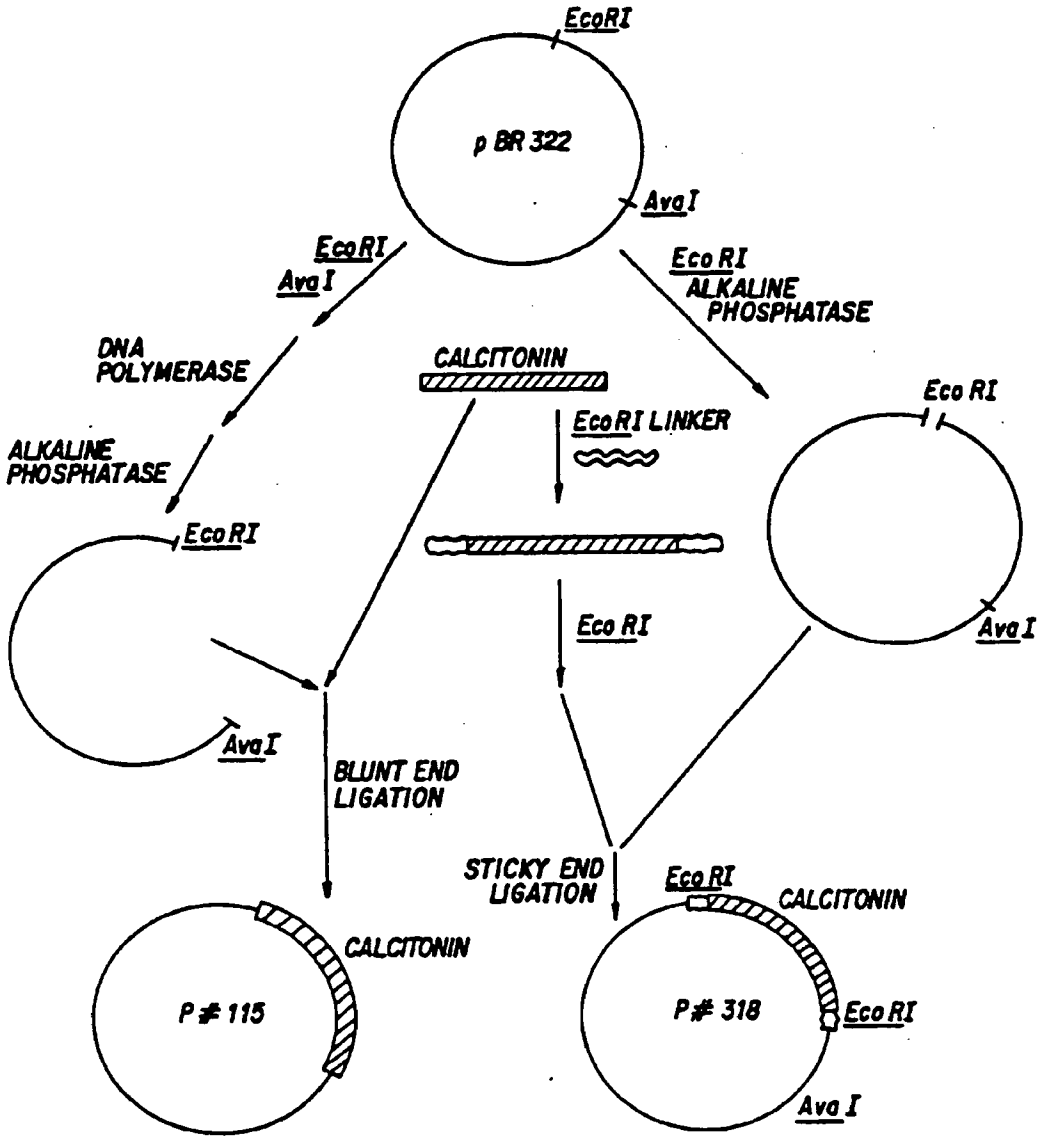
EcoRI
p BR 322
AvaI
EcoRI
AvaI
Eco RI
ALKALINE
PHOSPHATASE
DNA
POLYMERASE
CALCITONIN
EcoRI LINKER
Eco RI
ALKALINE
PHOSPHATASE
EcoRI
AvaI
Eco RI
AvaI
BLUNT END
LIGATION
STICKY END
LIGATION
Eco RI
CALCITONIN
CALCITONIN
P≠ 115
P≠ 318
Eco RI
AvaI

FIG. 10

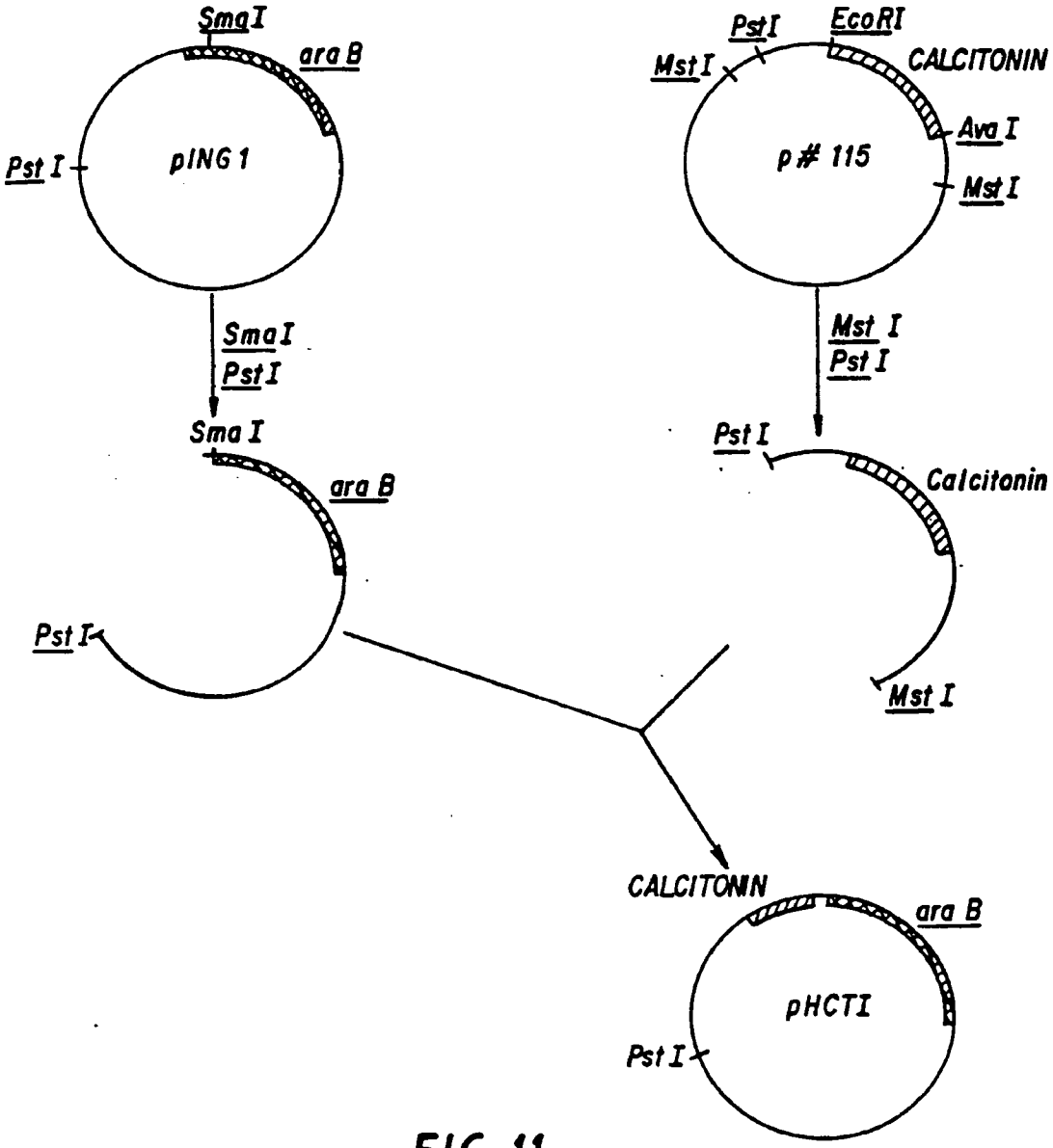
Sma I
ara B
Pst I
pING 1
Sma I
Pst I
Sma I
ara B
Pst I
Pst I
EcoRI
Mst I
CALCITONIN
Ava I
p # 115
Mst I
Mst I
Pst I
Pst I
Calcitonin
Mst I
CALCITONIN
ara B
pHCTI
Pst I

FIG. 11

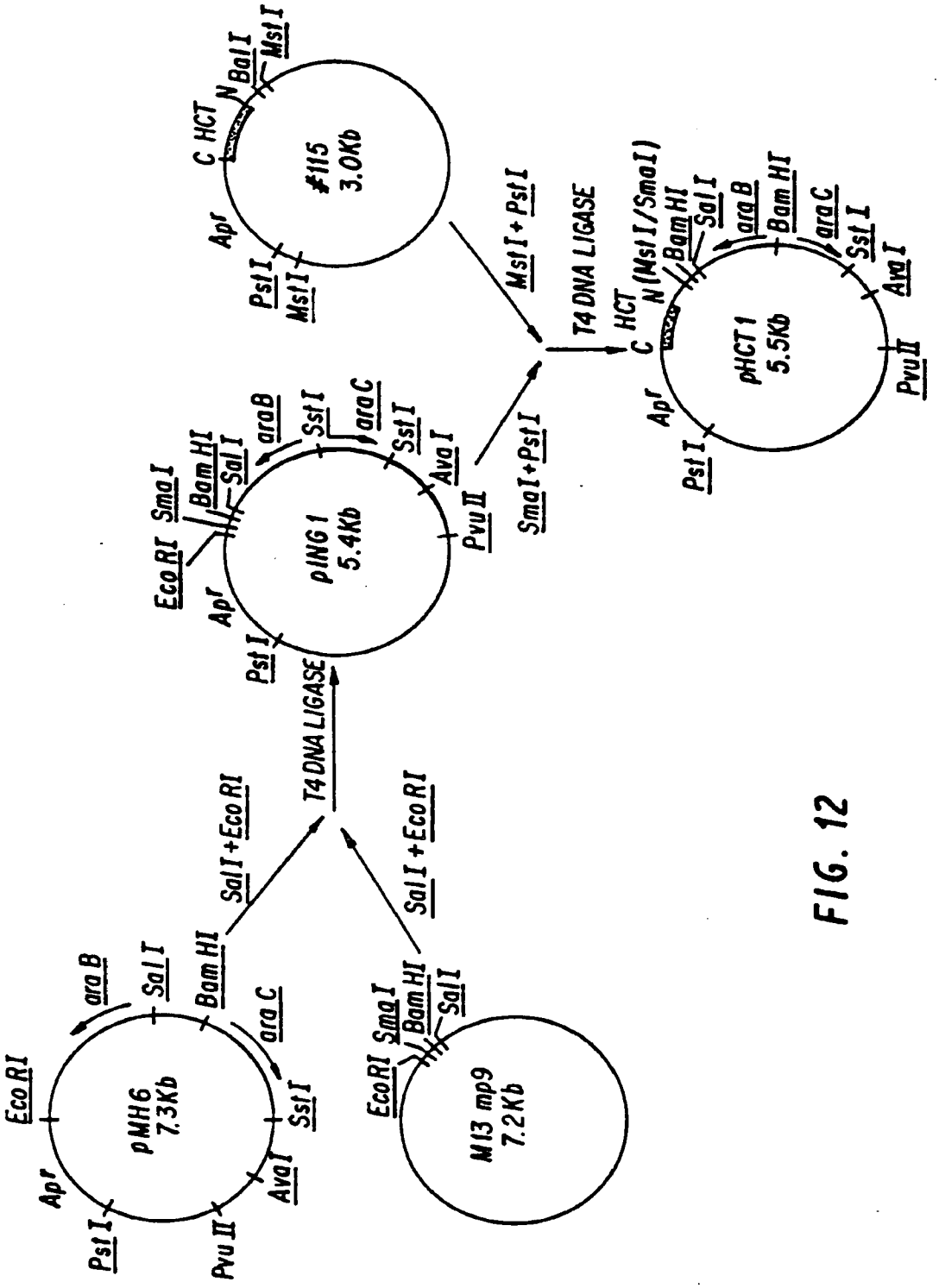
pMH6
7.3Kb
Eco RI
Ap^r
Pst I
Pvu II
Ava I
Sst I
ara C
Bam HI
Sal I
ara B
Sal I + Eco RI
T4 DNA LIGASE
M13 mp9
7.2Kb
Eco RI
Sma I
Bam HI
Sal I
pING 1
5.4Kb
Ap^r
Pst I
Eco RI
Sma I
Bam HI
Sal I
ara B
Sst I
ara C
Sst I
Ava I
Pvu II
Sma I + Pst I
T4 DNA LIGASE
φ115
3.0Kb
Ap^r
Pst I
Msl I
C HCT N
Bal I
Msl I
Msl I + Pst I
pHCT1
5.5Kb
Ap^r
Pst I
C HCT N (Msl I / Sma I)
Bam HI
Sal I
ara B
Bam HI
ara C
Sst I
Ava I
Pvu II

FIG. 12